

161444

Shears, Beverly

From: Devi, Sarvamangala  
Sent: Wednesday, July 27, 2005 3:59 PM  
To: Shears, Beverly  
Subject: 10/070,882

Beverly:

In application 10/070,882, please perform a sequence search for SEQ ID NO: 2 in commercial and pending databases.

Thanx.

S. DEVI, Ph.D.  
Primary Examiner  
AU 1645  
Rems - 3C18

UA  
Need pending

1

Date completed:

Searcher:

Beverly e 2528

Terminal time:

Elapsed time:

CPU time:

Total time:

Number of Searches:

Number of Databases:

Search Site

STIC

CM-1

Pre-S

Type of Search

N.A. Sequence

A.A. Sequence

Structure

Bibliographic

Vendors

IG

STN

Dialog

APS

Geninfo

SDC

DARC/Questel

Other

CGN

**This Page Blank (uspto)**



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number 161444**

**TO: Sarvamangala Devi**  
**Art Unit: 1645**  
**Location: REM 3C18**  
**Serial Number: 10/070882**

**Tuesday, August 09, 2005**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**REM 1A54**  
**Phone: 571-272-2528**  
**beverly.shears@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

Off nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:30:55 ; Search time 7070 Seconds  
(without alignments)  
748.364 Million cell updates/sec

Title: US-10-070-882A-2

Perfect score: 139  
Sequence: 1 gtgactctgtcgcacgaact.....ctgtcttctatgttaacaca 139

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_hlc:\*  
4: gb\_esc3:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	34.8	25.0	3932	3	AK082454 Mus muscu
2	34.4	24.7	414	6	BY642267 BY642267
3	33.6	24.2	204	1	AV231406 AV231406
4	33.6	24.2	274	2	BB775700 BB775700
5	33.6	24.2	318	2	BB094276 BB094276
6	33.6	24.2	335	2	BB222962 BB222962
7	33.6	24.2	448	5	BY439440 BY439440
8	33.6	24.2	570	8	AZ435636 IM0222H14
9	33.6	24.2	590	2	BE11381 I48-1355T
10	33.6	24.2	605	4	BM658097 MCR059H01
11	33.6	24.2	625	6	CA835622 MCS039B06
12	33.6	24.2	633	6	CA834105 MCS028H02
13	33.6	24.2	674	4	BM301497 MCR046F12
14	33.6	24.2	699	4	BM301422 MCR045F12
15	33.6	24.2	746	9	CR303205 MCR045F12
16	33.6	24.2	863	3	BE035174 MCR02B09
17	33.6	24.2	866	9	CR308064 MCR02B09
18	33.6	24.2	889	9	CG952370 MCR02B09
19	33.6	24.2	1192	2	BE037111 MCR02B09
20	33.4	24.0	521	8	AQ0830244 HS 4825_B
21	33.4	24.0	728	9	AG308530 Mus muscu
22	33.4	24.0	957	6	CA791309 AGENCOURT
23	33.2	23.9	205	2	BB004186 BB004186
24	33	23.7	467	8	B2915818 B2915818

25	33.8	23.6	206	1	AV336201 AV336201
26	33.8	23.6	690	8	BZ035953 BZ035953
27	33.4	23.3	252	2	BB230680 BB230680
28	33.4	23.3	270	2	BB118576 BB118576
29	32.4	23.3	533	5	BY480819 BY480819
30	32.4	23.3	906	8	AQ781019 HS_3169_B
31	32.2	23.2	275	1	AI414763 ma59d02.x
32	32.2	23.2	459	2	BE289932 601089159
33	32.2	23.2	579	1	AI981576 Pat. PK006
34	32.2	23.2	741	9	CC537179 CC537179
35	32.2	23.2	769	8	AQ869785 AQ869785
36	32	23.0	239	1	AA686653 AA686653
37	32	23.0	274	2	BB364524 BB364524
38	32	23.0	320	2	BB464866 BB464866
39	32	23.0	391	5	BY468249 BY468249
40	32	23.0	490	8	AQ430313 HS_5093_A
41	32	23.0	501	8	BH002081 BH002081
42	32	23.0	611	7	CO040889 CO040889
43	32	23.0	706	8	BH830539 BH830539
44	32	23.0	745	9	AG479014 Mus muscu
45	31.8	22.9	175	9	CE320783 CE320783

## ALIGNMENTS

RESULT 1  
AK082454/c 3932 bp mRNA linear HTC 03-APR-2004  
LOCUS  
DEFINITION Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone: C230052F14 product: unclassified, full  
insert sequence.

ACCESSION AK082454  
VERSION AK082454.1 GI:26100683  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Carninci, P., and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636

REFERENCE  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159

REFERENCE  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tahito, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, H., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multichannel sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861

REFERENCE  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

TITLE	JOURNAL	REFERENCE	AUTHORS
Group Phase I & II Team.			
Analysis of the mouse transcriptome based on functional annotation			
Of 60,770 full-length cDNAs			
Nature 420, 563-573 (2002)			
(bases 1 to 3932)			
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagiri, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirokane, T., Hoti, F., Imomani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kohjiya, Y., Kondo, S., Komori, H., Koude, M., Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Ohashi, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Yamatsuta, M., and Hayashizaki, Y.			
Muramatsu, M. and Hayashizaki, Y.			
Direct Submission			
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Shinhito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-research@riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.			
Please visit our web site for further details.			
URL: http://genome.gsc.riken.jp/			
URL: http://fantom.gsc.riken.jp/			
Location/Qualifiers			
1..3932			
/organism="Mus musculus"			
/mol_type="mRNA"			
/blast_type="C57BL/6J"			
/db_xref="PANTOM_DB:C230052F14"			
/db_xref="taxon:10090"			
/clone="C230052F14"			
/issue_type="cerebellum"			
/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="0 day neonate"			
1..3932			
/note="unclassifiable"			
misc_feature			
Query Match	25.0%	Score 34.8;	DB 3; Length 3932;
Best Local Similarity	54.8%;	Pid.No. 13;	
Matches	69;	Conservative	0; Mismatch 57; Indels 0; Gaps 0;
OY	14	ACGAAGTAATTAATGCGCTGCCTCCCTCTTTTCTTCAGAAAGAGGTACTATTGTC	73
Db	2194	ACTATCTCATTTAAAGCTCTTCACTCTGTGTATATAAAGTTTGTTGTGGATTTTG	2135
OY	74	TGGTTATTAACTGTTTATGCCCAAGCACATTAACGCTGACTGTTTATTGTT	133
Db	2134	TGGTATATGAAGTGAATTTTTCAGAAATCAATGAATTAAGTTGATTAGCAAGTGT	2075
OY	134	AACACA	139
Db	2074	CAAGCA	2069
RESULT 2			
Bt642267			
LOCUS	BY642267	414 bp	mRNA linear EST 15-DEC-2002
DEFINITION	BN642267 RIKEN full-length enriched, visual cortex Mus musculus		
ACCESION	BY642267		
VERSION	BY642267.1	GI:26977449	
KEYWORDS	EST.		

**SOURCE**  
 Mus musculus  
 Mus musculus (house mouse)

**ORGANISM**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathae; Muridae; Murinae; Mus.

**REFERENCE**  
 1 (bases 1 to 414)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishikado, I., Otsu, N., Saito, R., Suzuki, H., Yamakawa, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Kiyosawa, S., Beisel, K. W., Blake, J. A., Brad, D., Brunic, V., Chotia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragan, T. A., Fletcher, C. F., Forrest, A., Frazier, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okimoto, T., Pavan, W. M., Pettes, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. V., Qi, D., Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Rindwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Walstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Cinnici, P., Hayatsu, N., Hirozane-Kobayashi, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aekawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kasawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, R. and Hayashizaki, Y.

**TITLE**  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL**  
 Nature 420, 563-573 (2002)

**MEDLINE**  
 22354683

**PUBMED**  
 12466851

**COMMENT**  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Shuho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
 Aizawa, K., Akimura, T., Aekawa, T., Carrinci, P., Fukuda, S., Hirozane, A., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numasaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Tissues were provided by Michela Fegiolini and Takao K. Henrich (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES  
source  
location/Qualifiers  
1..414  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K430350A22"  
/issue\_type="visual cortex"  
/clone\_1ib="RIKEN full-length enriched, visual cortex"

ORIGIN  
Query Match 24.7%; Score 34.4; DB 6; Length 414;  
Best Local Similarity 57.4%; Pred. No. 12; Mismatches 46; Indels 0; Gaps 0;  
Matches 62; Conservative 0;

Qy 31 CTGCTCAACCTCTTTCTTTCAGAAAGGGTGACTATTGTCGTTTATTACTGTTT 90  
214 CTGCTCAACCTCTTTCTTTCAGAAAGGGTGACTATTGTCGTTTATTACTGTTT 273

Db 91 ATCCCAAGACACCATATCAACGCTAGACTGTTCTTATTGTTAAC 138  
274 TTTCTCAGACACACCCACATTCGTTGACGTCCTGTTGTCAGCGC 321

RESULT 3  
AV231406 204 bp mRNA linear EST 15-NOV-2001  
AV231406 RIKEN full-length enriched, 0 day neonate skin Mus  
musculus cDNA clone 4632406L08 3', mRNA sequence.  
AV231406.1 GI:6183921

ACCESSION  
AV231406.1 GI:6183921

VERSION  
EST.

KEYWORDS  
Mus musculus (house mouse)

SOURCE  
Mus musculus

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 204)

REFERENCE  
AUTHORS  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,  
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F.,  
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,  
Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,  
Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suganara, Y.,  
Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomimaga, N.,  
Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A.,  
Yokota, Y., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN House ESTs (Konno, H., et al. 1999)  
Unpublished (1999)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.jp, URL: http://genome-gsc.riken.jp/  
Sasaki, M., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,  
Matsuyama, S., Carninci, P., Muramatsu, M., Okazaki, Y. and  
Hayashizaki, Y.  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome-gsc.riken.go.jp) for  
further details.  
Location/Qualifiers

source  
1..204  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="4632406L08"  
/sex="mixed"  
/issue\_type="skin"  
/dev\_stage="0 day neonate"  
/lab\_host="DH10B"  
/clone\_1ib="RIKEN full-length enriched, 0 day neonate  
skin"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGGAGAGAGATCCAGACCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 10.0 and subtraction to Rot = 100.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGGAGAGATTCGATTAATTAATTAATCCCCCCCCC  
3']. cDNA was cloned into the XhoI and BamHI sites.  
Vector: a modified plasmid KS(+) after bulk excision  
from Lambda FLC 1"

ORIGIN  
Query Match 24.2%; Score 33.6; DB 1; Length 204;  
Best Local Similarity 57.7%; Pred. No. 18; Mismatches 44; Indels 0; Gaps 0;  
Matches 60; Conservative 0;

Qy 31 CTGCTCAACCTCTTTCTTTCAGAAAGGGTGACTATTGTCGTTTATTACTGTTT 90  
10 CTGCTCAACCTCTTTCTTTCAGAAAGGGTGACTATTGTCGTTTATTACTGTTT 69

Db 91 ATCCCAAGACACCATATCAACGCTAGACTGTTCTTATTGTTAAC 134  
70 TTTCTCAGACACACCCACATTCGTTGACGTCCTGTTGTCAGCGC 113

RESULT 4  
BB775700 274 bp mRNA linear EST 08-JUL-2003  
BB775700 RIKEN full-length enriched, RCB-0559 K-1. F1 cDNA Mus  
musculus cDNA clone 4632406L09 3', mRNA sequence.  
BB775700.1 GI:16936380

ACCESSION  
BB775700.1 GI:16936380

VERSION  
EST.

KEYWORDS  
Mus musculus (house mouse)

SOURCE  
Mus musculus

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 274)

REFERENCE  
AUTHORS  
Akimura, T., Arikawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hirozane, T., Hirozane, T., Imotani, K.,  
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,  
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Kakita, S.,  
Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)

1-7-22	Sheniro-cho,	Tsurumi-ku,	Yokohama,	Kanagawa	230-0045,	Japan
Tel:	81-45-503-9222					
Fax:	81-45-503-9216					
Email:	genome-res@isc.riken.jp,	URL:	http://genome.gsc.riken.jp/			
Carninci, P.,	Shibata, Y.,	Hayatsu, N.,	Sugahara, Y.,	Shibata, K.,		
Itoh, M.,	Komoto, H.,	Okazaki, Y.,	Muramatsu, M.,	and Hayashizaki, Y.		
Normalization and subtraction of cap-trapper-selected cDNAs to						
prepare full-length cDNA libraries for rapid discovery of new						
genes. Genome Res.	10 (10),	1617-1630	(2000)			
wasgi, K.,	Fujiwara, S.,	Inoue, K.,	Togawa, Y.,	Izawa, M.,	Ohara, E.,	
Matsubara, S.,	Kawai, J.,	Okazaki, Y.,	Ishikawa, T.,	Osawa, K.,	Tanaka, T.,	
and Hayashizaki, Y.						
RIKEN integrated sequence analysis (RISA) system-384-format						
sequencing pipeline with 384 multicapillary sequencer. Genome Res.						
10 (11),	1757-1771	(2000)				
Komoto, H.,	Fukunishi, Y.,	Shibata, K.,	Itoh, M.,	Carninci, P.,		
Sugahara, Y.,	and Hayashizaki, Y.					
Computer-based methods for the mouse full-length cDNA						
encyclopedic real-time sequence clustering for construction of a						
nonredundant cDNA library. Genome Res.	11 (2),	281-289	(2001)			
Please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for						
further details.						
e mouse tissues.						
location/Qualifiers						
1..274						
/organism="Mus musculus"						
/mol_type="mRNA"						
/strain="DDD"						
/db_xref="taxon:10090"						
/clone="G430026109"						
/cell_line="RCB-0559 K-1.fl"						
/clone_idb="RIKEN full-length enriched, RCB-0559 K-1 . fl						
cDNA"						
/note="Organ: spleen"						
ORIGIN						
Query Match	24.2%	Score	33.6;	DB 2;	Length	274;
Best Local Similarity	57.7%	Pred. No.	19;			
Matches	60;	Conservative	0;	Mismatches	44;	Indels 0;
Db						
31	CTGCTCAACCCCTTTTCTTCAAGAAAGGGGACTATTGTCGTGTTATTAACGTGT	90				
80	CTGCTCCAGCTCTTCTCTTCTTGACCTGTCGTCCACTTTTCAGATCTTTCACGT	139				
Qy	91	ATCCCAAGACCATTAATCAACGCTAGACTGTTCTTATGTTA	134			
Db	140	TTCTTCAGAGACACCCACACATTTGTTACAGTCTTTTTC	183			
RESULT 5						
BB094276						
LOCUS						
DEFINITION	BB094276	318 bp	mRNA	linear	EST 26-JUN-2000	
ACCESSION	BB094276					
VERSION	BB094276.1					
KEYWORDS	EST.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
REFERENCE	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 318)					
	Komoto, H.,	Aizawa, K.,	Akaiwa, S.,	Akiyama, J.,	Arakawa, T.,	
	Carninci, P.,	Endo, T.,	Fukuda, S.,	Fukunishi, Y.,	Hata, A.,	Hayatsu, N.,
	Hirozane, T.,	Hori, F.,	Ishii, Y.,	Ishikawa, J.,	Ishikawa, T.,	Itoh, M.,
	Izawa, M.,	Kadota, K.,	Kagawa, I.,	Kai, C.,	Kawai, J.,	Kikuchi, N.,
	Kiyosawa, H.,	Kojima, Y.,	Kondo, S.,	Koya, S.,	Kurihara, C.,	
	Kusabe, M.,	Matsuyama, T.,	Miki, R.,	Mizuno, Y.,	Nakamura, M.,	Oda, H.,
	Okazaki, Y.,	Ono, T.,	Owa, C.,	Saito, H.,	Sakai, C.,	Sato, K.,
	Shibata, K.,	Shibata, Y.,	Shinmoto, Y.,	Shinagawa, A.,	Shitaki, T.,	
	Sogabe, Y.,	Sugahara, Y.,	Suuki, H.,	Suzuki, H.,	Tagawa, A.,	

TITLE  
JOURNAL  
COMMENT

RIKEN Mouse ES-Str (Korn, H., et al.)  
Unpublished (2000)  
Contact: Yoshinori Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,  
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,  
Thermotranscription and thermotranscription of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Kikunishi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,  
Okazaki, Y., and Hayashizaki, Y.,  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.,  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.riken.go.jp>) for  
further details.

Location/Qualifiers

1. 318  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="9430045021"  
/tissue\_type="embryonic body between diaphragm region and  
neck"  
/dev\_stage="12 days embryo"  
/lab\_host="DH10B"  
/clone\_1db="RIKEN full-length enriched, 12 days embryo,  
embryonic body between diaphragm region and neck"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGGAGAGAGAGATCCAGACGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 20.0 and subtraction to Rot = 370.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGGAGAGATTCGAGGTATTAATTAATTCACCCCCCCCC  
3']. cDNA was cleaved with XhoI and BamHI."

ORIGIN

Query Match 24.2%; Score 33.6; DB 2; Length 318;  
Best Local Similarity 57.7%; Pred. No. 20;  
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 31 CTGCTCAACCTCTTTTCTTCAGAAAGAGGTGACATTTGTCGTGTTATTAACCTGTT 90  
|||||  
DB 122 CTGCTCCAGAGATCTTTCTTTCTTCAGCGTGTGCTCCCACTGTTGAGATTTTTCACCTGTT 181  
|||||

QY 91 ATCCCAAGACCATTAATCAACCTGACGTGTTCTTATGTTA 134  
|||||

DB 182 TTCTTCAGACACCCCAACATTCCTGACAGCTCTTGTTGTC 225  
|||||

RESULT 6  
BB222962



Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct SubMISSION Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

# FEATURES

## source

```
1..448
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1920162D12"
/clone_lib="RIKEN full-length enriched, pooled tissues, 16 days embryo, etc."
/notes="pooled tissues: (dev_stage=16 days embryo, tissue_type=heart, sex=mix), (dev_stage=16 days embryo, tissue_type=kidney, sex=mix), (dev_stage=17 days embryo, tissue_type=heart, sex=mix), (dev_stage=17 days embryo, tissue_type=stomach, sex=mix), (dev_stage=17 days embryo, tissue_type=kidney, sex=mix), (dev_stage=17 days pregnant, adult, tissue_type=salmon, sex=female), (dev_stage=13 days embryo, tissue_type=liver, sex=mix)"
```

# ORIGIN

Query Match 24.2%; Score 33.6; DB 5; Length 448;  
 Best Local Similarity 61.4%; Pred. No. 21;  
 Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 47 TCTTCAGAAAGAGGCTGACTATTTGTCGTTTATTACTGTTATCCCAAGACCAT 106  
 |||||  
 DB 262 TCTTCGAGCTGTGGGTCCGCTTCAGATCTCTTCACTGTTTCTTCAGAGCACC 321  
 |||||  
 QY 107 AATCAAGCTAGACTGTTCTTATTGTTA 134  
 |||||  
 DB 322 ACACATGCTTGACAGCTCTTGTGCA 349  
 |||||

RESULT 8  
 AZ435636 570 bp DNA linear GSS 03-OCT-2000  
 LOCUS 1M0222H14R Mouse 10kb plasmid UGCGIM library Mus musculus genomic  
 DEFINITION clone UGCGIM0222H14 R, genomic survey sequence.  
 ACCESSION AZ435636  
 VERSION AZ435636.1 GI:10559649  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 570)  
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamli, C., Islem, H., Longacre, S., Mamoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.  
 Mouse whole genome scaffolding with paired end reads from 10kb

# JOURNAL COMMENT

plasmid inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0222 row: H column: 14  
 Seq primer: CACACGAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 570.

# FEATURES

## source

```
1..570
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCGIM0222H14"
/sex="Male"
/lab_host="B. CoII strain X110-Gold, T1-resistant, F-"  

/clone_lib="Mouse 10kb plasmid UGCGIM library"  

/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli X110-Gold (Stratagene) cells and selected for ampicillin resistance."
```

# ORIGIN

Query Match 24.2%; Score 33.6; DB 8; Length 570;  
 Best Local Similarity 63.8%; Pred. No. 22;  
 Matches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 GTGACTCTGTGAGCAACTTAATATGCGCTGACCTCTTTCTTTCAGAAAGAG 60  
 |||||  
 DB 216 GTGACTTATGTAAGCTATTAATTTCTCCATTCTCATGCTTACTTATTAAC 157  
 |||||  
 QY 61 GTGACTATTTGTCGTGTTA 80  
 |||||  
 DB 156 GAGAAATGTTTGTTCCTTA 137  
 |||||

RESULT 9  
 BE131381 590 bp mRNA linear EST 20-FEB-2001  
 LOCUS 148-1355T3 Ice plant lambda Uni-Zap XR expression library, 48 hours  
 DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone 148-1355,  
 mRNA sequence.  
 ACCESSION BE131381  
 VERSION BE131381.1 GI:8578744  
 KEYWORDS EST.  
 SOURCE Mesembryanthemum crystallinum (common iceplant)  
 ORGANISM Mesembryanthemum crystallinum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Alzooceae; Mesembryanthemum.



REFERENCE 1 (bases 1 to 590)  
AUTHORS Cushman, J.C.  
TITLE An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
JOURNAL Unpublished (1997)  
COMMENT Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers  
FORWARD: T7  
BACKWARD: T3  
Plate: 148-14 row: E column: 7  
Seq primer: T3  
High quality sequence stop: 350  
POLA-No.  
FEATURES  
source location/Qualifiers  
1..590  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="148-1355"  
/tissue\_type="leaf, 48 h 0.4M NaCl"  
/dev\_stage="Six week old"  
/clone\_lib="Ice plant lambda Uni-Zap XR expression  
library, 48 hours NaCl treatment"  
/note="Vector: lambda Uni-Zap XR, Bluescript SK-, Site\_1:  
EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 24.2%; Score 33.6; DB 2; Length 590;  
Best Local Similarity 61.4%; Pred. No. 22;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 44 TTTTCTTGAGAAAGAGGACTATTGTCGTGTTTAACTGTTATCCCAAGCAC 103  
|||||  
147 TTTTCATCAAAAGCAGTGACTTGGTTGAATTTCTATCCCTAATCAACAATCAC 206  
|||||  
QY 104 CATATCAACGCTAGACTGTTCTTATG 131  
|||||  
DB 207 CTGAAAAAACAGTGGAGTTGTTTGG 234  
|||||

RESULT 10 605 bp mRNA linear EST 26-FEB-2002  
BM658097  
LOCUS MCR059H01.69977 Ice plant Lambda Uni-Zap XR expression library, 48  
hours NaCl treatment prescreened for removal of highly abundant  
transcripts Mesembryanthemum crystallinum cDNA clone MCR059H01.5,  
mRNA sequence.  
BM658097  
VERSION BM658097.1 GI:18957608  
KEYWORDS EST.  
SOURCE Mesembryanthemum crystallinum (common iceplant)  
ORGANISM Mesembryanthemum crystallinum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Aizoaceae; Mesembryanthemum.  
1 (bases 1 to 605)  
Cushman, J.C.  
An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
Unpublished (1997)  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

PCR Primers  
FORWARD: T3 20mer  
BACKWARD: T7 21mer  
Plate: 059 row: H column: 01  
Seq primer: T3 20mer  
High quality sequence stop: 605.  
FEATURES  
source location/Qualifiers  
1..605  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="MCR059H01"  
/tissue\_type="leaf"  
/dev\_stage="six-week-old"  
/clone\_lib="Ice plant lambda Uni-Zap XR expression  
library, 48 hours NaCl treatment prescreened for removal  
of highly abundant transcripts"  
/note="Vector: lambda Uni-Zap XR, Bluescript SK-, Site\_1:  
EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 24.2%; Score 33.6; DB 4; Length 605;  
Best Local Similarity 61.4%; Pred. No. 22;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 44 TTTTCTTGAGAAAGAGGACTATTGTCGTGTTTAACTGTTATCCCAAGCAC 103  
|||||  
153 TTTTCATCAAAAGCAGTGACTTGGTTGAATTTCTATCCCTAATCAACAATCAC 212  
|||||  
QY 104 CATATCAACGCTAGACTGTTCTTATG 131  
|||||  
DB 213 CTGAAAAAACAGTGGAGTTGTTTGG 240  
|||||

RESULT 11 625 bp mRNA linear EST 12-DEC-2002  
CA835622  
LOCUS MCS039B06.160752 Ice plant lambda Uni-Zap XR expression library, 5  
days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase 1 (2  
AM). Mesembryanthemum crystallinum cDNA clone MCS039B06.5, mRNA  
sequence.  
CA835622  
VERSION CA835622.1 GI:26563387  
KEYWORDS EST.  
SOURCE Mesembryanthemum crystallinum (common iceplant)  
ORGANISM Mesembryanthemum crystallinum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Aizoaceae; Mesembryanthemum.  
1 (bases 1 to 625)  
Cushman, J.C.  
An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
Unpublished (1997)  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers  
FORWARD: T3 20mer  
BACKWARD: T7 21mer  
Plate: 039 row: B column: 06  
Seq primer: T3 20mer  
High quality sequence stop: 625.  
FEATURES  
source location/Qualifiers  
1..625  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="MCS039B06"

/tissue\_type="leaf"  
/dev\_stage="five-week-old"  
/clone\_lib="Ice plant lambda Uni-Zap XR expression  
library", 5 days 0.5 M NaCl treatment, Crassulacean acid  
metabolism, phase I (2 AM)."  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site\_1:  
EcoRI; Site\_2: XhoI; library construction was performed  
according to Stratagene's recommended protocol for the  
Lambda UniZapXR vector and cDNA synthesis kit."

## ORIGIN

Query Match 24.2%; Score 33.6; DB 6; Length 625;  
Best Local Similarity 61.4%; Pred. No. 22;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 44 TTTTCTTCAAGAAAGGAGTACTATTGTCGTTTAACTGTTATCCCAAGAC 103  
DB 156 TTTTCATCACAAGAGGAGTGTGGTTGAAATTTCTATCCCTAATCAACAATC 215  
QY 104 CATATCAACGCTAGACTGTTCTTATG 131  
DB 216 CTGAATAAAACAGTGGAGTTTGTG 243

RESULT 12  
CA834105 633 bp mRNA linear EST 12-DEC-2002  
LOCUS MCS028H02.153488 Ice plant lambda Uni-Zap XR expression library, 5  
DEFINITION days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2  
AM). Mesembryanthemum crystallinum cDNA clone MCS028H02.5, mRNA  
sequence.

ACCESSION  
CA834105  
VERSION  
CA834105.1 GI:26561870

KEYWORDS  
SOURCE  
ORGANISM  
Mesembryanthemum crystallinum (common iceplant)

REFERENCE  
AUTHORS  
TITLE  
Cushman, J.C.  
An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
Unpublished (1997)

JOURNAL  
COMMENT  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu

PCR Primers  
FORWARD: T3 20mer  
BACKWARD: T7 21mer  
Plate: 028 row: H column: 02  
Seq primer: T3 20mer  
High quality sequence stop: 633.  
Location/Qualifiers  
1..633

FEATURES  
source  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="MCS028H02"  
/tissue\_type="leaf"  
/dev\_stage="five-week-old"  
/clone\_lib="Ice plant lambda Uni-Zap XR expression  
library, 5 days 0.5 M NaCl treatment, Crassulacean acid  
metabolism, phase I (2 AM)."  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site\_1:  
EcoRI; Site\_2: XhoI; library construction was performed  
according to Stratagene's recommended protocol for the  
Lambda UniZapXR vector and cDNA synthesis kit."

ORIGIN

Query Match 24.2%; Score 33.6; DB 6; Length 633;  
Best Local Similarity 61.4%; Pred. No. 22;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 44 TTTTCTTCAAGAAAGGAGTACTATTGTCGTTTAACTGTTATCCCAAGAC 103  
DB 165 TTTTCATCACAAGAGGAGTGTGGTTGAAATTTCTATCCCTAATCAACAATC 224  
QY 104 CATATCAACGCTAGACTGTTCTTATG 131  
DB 225 CTGAATAAAACAGTGGAGTTTGTG 252

RESULT 13  
BM301497 674 bp mRNA linear EST 22-JAN-2002  
LOCUS MCR046F12.26036 Ice plant lambda Uni-Zap XR expression library, 48  
DEFINITION hours NaCl treatment prescreened for removal of highly abundant  
transcripts Mesembryanthemum crystallinum cDNA clone MCR046F12.5,  
mRNA sequence.

ACCESSION  
BM301497  
VERSION  
BM301497.1 GI:18023872

KEYWORDS  
SOURCE  
ORGANISM  
Mesembryanthemum crystallinum (common iceplant)

REFERENCE  
AUTHORS  
TITLE  
Cushman, J.C.  
An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
Unpublished (1997)

JOURNAL  
COMMENT  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu

PCR Primers  
FORWARD: T3 20mer  
BACKWARD: T7 21mer  
Plate: 046 row: F column: 12  
Seq primer: T3 20mer  
High quality sequence stop: 674.  
Location/Qualifiers  
1..674

FEATURES  
source  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="MCR046F12"  
/tissue\_type="leaf"  
/dev\_stage="six-week-old"  
/clone\_lib="Ice plant lambda Uni-Zap XR expression  
library, 48 hours NaCl treatment prescreened for removal  
of highly abundant transcripts"  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site\_1:  
EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 24.2%; Score 33.6; DB 4; Length 674;  
Best Local Similarity 61.4%; Pred. No. 22;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 44 TTTTCTTCAAGAAAGGAGTACTATTGTCGTTTAACTGTTATCCCAAGAC 103  
DB 156 TTTTCATCACAAGAGGAGTGTGGTTGAAATTTCTATCCCTAATCAACAATC 215  
QY 104 CATATCAACGCTAGACTGTTCTTATG 131  
DB 216 CTGAATAAAACAGTGGAGTTTGTG 243



RESULT 14  
BM301422 699 bp mRNA linear EST 22-JAN-2002  
LOCUS MCR045f12.25886 Ice plant Lambda Uni-Zap XR expression library, 48  
DEFINITION hours NaCl treatment prescreened for removal of highly abundant  
transcripts Mesembryanthemum crystallinum cDNA clone MCR045f12.5,  
mRNA sequence.  
ACCESSION BM301422  
VERSION BM301422.1 GI:18023797  
KEYWORDS  
SOURCE Mesembryanthemum crystallinum (common iceplant)  
ORGANISM Mesembryanthemum crystallinum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Aizoaceae; Mesembryanthemum.  
1 (bases 1 to 699)  
Cushman, J.C.  
REFERENCE An expressed sequence tag database for the common ice plant,  
AUTHORS Mesembryanthemum crystallinum  
TITLE Mesembryanthemum crystallinum  
JOURNAL Unpublished (1997)  
COMMENT Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers  
FORWARD: T3 20mer  
BACKWARD: T7 21mer  
Plate: 045 row: F column: 12  
Seq primer: T3 20mer  
High quality sequence stop: 699.  
Location/Qualifiers  
1..699  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="MCR045f12"  
/issue\_type="leaf"  
/dev\_stage="six-week-old"  
/clone\_1lb="Ice plant Lambda Uni-Zap XR expression  
library, 48 hours NaCl treatment prescreened for removal  
of highly abundant transcripts"  
/note="Vector: lambda Uni-Zap XR, Bluescript SK-, Site\_1:  
EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 24.2%; Score 33.6; DB 4; Length 699;  
Best Local Similarity 61.4%; Pred. No. 23;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 44 TTTTCTTGAGAAAGAGCGTACTATTGCTGTTTATTAACGTTTATCCCAAGCAG 103  
|||||  
DB 156 TTTTCATCAGAAAGCGAGTGAATTCCTTAATCCCTAATCAACAATCAG 215  
|||||  
QY 104 CATATCAAGCTAGACTGTTCTTATTTG 131  
|||||  
DB 216 CTGAAAAAAGGAGGAGTTGTTTG 243  
|||||

RESULT 15  
LOCUS CR303205 746 bp DNA linear GSS 28-FEB-2004  
DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago  
truncatula, genomic survey sequence.  
ACCESSION CR303205  
VERSION CR303205.1 GI:44708225  
KEYWORDS GSS.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 746)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage;  
BP 191 91006 Evry cedex - FRANCE (E-mail: sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
FEATURES  
source  
1..746  
Location/Qualifiers  
/organism="Medicago truncatula"  
/mol\_type="genomic DNA"  
/cultivar="Jemalong A17"  
/db\_xref="taxon:3880"  
/clone\_1lb="MTE1"  
/note="Vector: pindigobac ; Site\_1: EcoRI ; Site\_2: EcoRI  
; Debelle F. and Chalhouh B.-Genoscope sequence ID :  
mtel-21p8FM1"

ORIGIN  
Query Match 24.2%; Score 33.6; DB 9; Length 746;  
Best Local Similarity 61.4%; Pred. No. 23;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 50 TCAGAAAGAGGCTACTATTGCTGCTTTATTAACGTTTATCCCAAGCACCATAAT 109  
|||||  
DB 517 TGATTAAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 576  
|||||  
QY 110 CAAGCTAGACTGTTCTTATTTGTATACA 137  
|||||  
DB 577 CAATGATATATCTATATCCAAATGAA 604  
|||||

Search completed: August 4, 2005, 01:47:13.  
Job time : 7078 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 20:47:40 ; Search time 276 Seconds  
(without alignment)  
2981.319 Million cell updates/sec

Title: US-10-070-882a-2

Perfect score: 139  
Sequence: 1 gtagctctgctgcacgaact.....ctgtctcatgttaacaca 139

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: N\_Geneseq\_16Dec04:\*  
1: geneeqn19808:\*  
2: geneeqn19908:\*  
3: geneeqn20008:\*  
4: geneeqn20018:\*  
5: geneeqn20028:\*  
6: geneeqn20038:\*  
7: geneeqn20048:\*  
8: geneeqn20058:\*  
9: geneeqn20068:\*  
10: geneeqn20078:\*  
11: geneeqn20088:\*  
12: geneeqn20098:\*  
13: geneeqn20108:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139	100.0	139	4	AAf82356
2	139	100.0	2802	12	ADG31136
3	40	28.8	1385	6	ABL57265
4	37.2	26.8	1385	6	ACA49482
5	32.6	23.5	266145	10	AD887477
6	31	22.3	772	6	ABQ72598
7	31	22.3	810	6	ABQ72685
8	30.8	22.2	19521	4	AAK81193
9	30.8	22.2	19521	8	ADA41637
10	30.8	22.2	19521	10	ADA57769
11	30.6	22.0	1539	8	ACA02032
12	30.6	22.0	2799	5	AAH67957
13	30.6	22.0	349980	5	AAH68533
14	30.4	21.9	31241	10	AAH63515
15	30.4	21.9	31241	10	ACC69145
16	30.4	21.9	31241	12	ADN48950
17	30.4	21.9	104644	6	ABQ99653
18	30.4	21.9	109559	13	ABD33157
19	30.2	21.7	110000	9	ADB12064_03
20	30.2	21.7	138363	13	ABD32624

21	30	21.6	622	12	ADJ38283	Adj38283 Plaetid d
22	30	21.6	10711	4	AAK74790	AAK74790 Human imm
23	29.8	21.4	301	3	AAH06526	AAH06526 Human imm
24	29.8	21.4	301	4	AAH93642	AAH93642 Human pro
25	29.8	21.4	301	4	AAH63734	AAH63734 Human pro
26	29.8	21.4	301	4	AAH02707	AAH02707 Prostate
27	29.8	21.4	301	4	AAH84956	AAH84956 Human pro
28	29.8	21.4	301	5	ACA59543	ACA59543 Prostate
29	29.8	21.4	301	6	AB195106	AB195106 Human P8D
30	29.8	21.4	301	8	ACC95270	ACC95270 Prostate
31	29.8	21.4	301	10	ADB13743	ADB13743 Human pro
32	29.8	21.4	301	10	ADG26159	ADG26159 Human pro
33	29.8	21.4	340	5	ABV07484	ABV07484 Human pro
34	29.8	21.4	409	5	ABV37417	ABV37417 Human pro
35	29.8	21.4	434	7	AD571710	AD571710 Human kid
36	29.8	21.4	434	7	AD571943	AD571943 Human kid
37	29.8	21.4	492	6	ABL64736	ABL64736 Lung can
38	29.8	21.4	492	6	ABL65652	ABL65652 Lung can
39	29.8	21.4	1140	10	ADC92238	ADC92238 E. faeciu
40	29.8	21.4	1448	2	AAZ33468	AAZ33468 Human pro
41	29.8	21.4	2143	2	AAZ42062	AAZ42062 Human end
42	29.8	21.4	2654	6	AB557323	AB557323 CDNA enco
43	29.8	21.4	2700	4	ABL28364	ABL28364 Drosophila
44	29.8	21.4	2840	12	ADP64619	ADP64619 DNA of hu
45	29.8	21.4	2962	10	ADB63189	ADB63189 Human CDN

## ALIGNMENTS

RESULT 1	AAf82356	standard; DNA; 139 BP.
ID	AAf82356	standard; DNA; 139 BP.
XX	AAf82356	
AC	AAf82356	
XX	AAf82356	
DT	22-JUN-2001	(first entry)
XX	22-JUN-2001	
DE	Salmonella typhimurium phoP gene promoter.	
XX	Salmonella typhimurium; phoP gene promoter; PphoP; antibacterial;	
KW	antiviral; vaccine; lacZ; pagC; outer membrane porin C; ompC;	
KW	transgenic microorganism; antigen production; antigen delivery;	
KW	infection; ds.	
XX	Salmonella typhimurium.	
OS	Salmonella typhimurium.	
XX	Salmonella typhimurium.	
PN	WO200119974-A2.	
XX	WO200119974-A2.	
XX	22-MAR-2001.	
XX	22-MAR-2001.	
PF	06-SEP-2000; 2000WO-GB003402.	
XX	06-SEP-2000; 2000WO-GB003402.	
PR	10-SEP-1999; 99GB-00021275.	
PR	12-JUL-2000; 2000GB-00017000.	
XX	12-JUL-2000; 2000GB-00017000.	
PA	(MINA) UK SEC FOR DEFENCE.	
XX	(MINA) UK SEC FOR DEFENCE.	
PI	Ttbbal1 RW, Bullifent HL;	
XX	Ttbbal1 RW, Bullifent HL;	
DR	WPI; 2001-328017/34.	
XX	WPI; 2001-328017/34.	
PT	New recombinant gut-colonizing microorganism, useful as vaccine	
PT	component, comprises construct containing phoP, pagC or ompC gene	
PT	promoter linked to nucleic acid encoding protein that induces immune	
PT	response against pathogen.	
XX	response against pathogen.	
PS	Claim 1, Fig 6, 33pp; English.	
XX	Claim 1, Fig 6, 33pp; English.	
XX	The present sequence is a DNA fragment comprising the phoP gene promoter.	
CC	DNA fragments comprising the Salmonella typhimurium phoP, pagC, and ompC	
CC	gene promoters were integrated into a vector used to transform a	
CC	recombinant gut-colonizing microorganism. The promoter was operably	

CC linked to a nucleic acid encoding a protein that is able to induce a  
 CC protective immune response against an organism in a mammal. The construct  
 CC is useful for enhancing expression of a desired protein at mucosal  
 CC effector sites. It is useful for delivering a variety of antigenic agents  
 CC which can be used to induce a protective immune response against a wide  
 CC range of pathogens such as *Bacillus anthracis*, *Bordetella pertussis*,  
 CC *Schistosoma mansoni*, herpes simplex virus, and *Mycobacterium tuberculosis*.  
 CC The three promoters (P(phop), P(pagC) and P(ompC)) are induced at  
 CC different stages in the infection process, and hence at different sites  
 CC in the body. This approach allows the induction of different immune  
 CC responses which provide protection against pathogens which colonise  
 CC different host cell compartments. The *Salmonella* vaccine vector system is  
 CC ideally suited to the delivery of many vaccine antigens since the vaccine  
 CC delivery mechanism accurately mimics the natural disease, entering the  
 CC body via the gut

XX Sequence 139 BP; 37 A; 33 C; 23 G; 46 T; 0 U; 0 Other;

Query Match 100.0%; Score 139; DB 4; Length 139;

Best Local Similarity 100.0%; Pred. No. 6.3e-36;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACTCTGTGTCGACGAACTTAATATAGCTGCTCACCCTCTTTCTTCGAAAGAG 60

DB 1 GTGACTCTGTGTCGACGAACTTAATATAGCTGCTCACCCTCTTTCTTCGAAAGAG 60

QY 61 GTGACTATTGTCTGTTTATTACTGTTATCCCAAGACCATATCAAGCTAGAC 120

DB 61 GTGACTATTGTCTGTTTATTACTGTTATCCCAAGACCATATCAAGCTAGAC 120

QY 121 TGTCTTATTGTTTAACACA 139

DB 121 TGTCTTATTGTTTAACACA 139

RESULT 2  
 ADG31136  
 ID ADG31136 standard; DNA; 2802 BP.

AC ADG31136;

DT 26-FEB-2004 (first entry)

DE *Salmonella typhimurium* PphoP and phoP operon DNA.

KW live attenuated derivative; pathogenic *Enterobacteriaceae*;

KW cross protective immunity; antibacterial; immunostimulant; vaccine;

OS *Salmonella typhimurium*.

PN WO2003096812-A1.

PD 27-NOV-2003.

PF 15-APR-2003; 2003WO-US011802.

PR 15-APR-2002; 2002US-0372616P.

PR 18-APR-2002; 2002US-0373626P.

PA (UNITW) UNIV WASHINGTON.

PI Curtiss R;

DR WPI; 2004-042484/04.

DR P-PSDB; ADG31137, ADG31138, ADG31139.

XX New live attenuated derivative of a pathogenic *Enterobacteriaceae*

PT species, useful as a vaccine for inducing cross protective immunity

PT against infections caused by various *Enterobacteriaceae* strains or

PT serotypes.

PS Example 17; Fig 28; 133pp; English.

XX The invention relates to a novel live attenuated derivative of a  
 CC pathogenic *Enterobacteriaceae* species having enhanced ability to induce  
 CC cross protective immunity against *Enterobacteriaceae*. The derivative of  
 CC the invention demonstrates antibacterial and immunostimulant activities  
 CC and may be useful as a vaccine for inducing a high level immune response  
 CC and/or cross protective immune response to protect individuals from  
 CC infection from a diversity of species or serotypes of bacterial  
 CC pathogens. The current sequence is that of the *Salmonella typhimurium*  
 CC PphoP and phoP operon DNA of the invention.

XX Sequence 2802 BP; 721 A; 702 C; 724 G; 655 T; 0 U; 0 Other;

Query Match 100.0%; Score 139; DB 12; Length 2802;

Best Local Similarity 100.0%; Pred. No. 1.5e-35;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACTCTGTGTCGACGAACTTAATATAGCTGCTCACCCTCTTTCTTCGAAAGAG 60

DB 511 GTGACTCTGTGTCGACGAACTTAATATAGCTGCTCACCCTCTTTCTTCGAAAGAG 570

QY 61 GTGACTATTGTCTGTTTATTACTGTTATCCCAAGACCATATCAAGCTAGAC 120

DB 571 GTGACTATTGTCTGTTTATTACTGTTATCCCAAGACCATATCAAGCTAGAC 630

QY 121 TGTCTTATTGTTTAACACA 139

DB 631 TGTCTTATTGTTTAACACA 649

RESULT 3  
 ABL57265  
 ID ABL57265 standard; DNA; 41 BP.

AC ABL57265;

DT 09-AUG-2002 (first entry)

DE *Escherichia coli* DNA 5' to phoA coding region.

KW Transcription terminator; trpA; attenuation; vaccine; virulence;

KW antibacterial; fungicide; antiparasitic; protozoacide; phoA; ds.

OS *Escherichia coli*.

FT Key Location/Qualifiers

FT misc\_feature 1..19

FT /\*tag= a

FT /note= "5' single-stranded overhang"

FT misc\_feature 41

FT /\*tag= b

FT /note= "5' overhang on complementary strand of 4 bases

FT with sequence 5'-GATC-3'"

PN WO200230457-A2.

PD 18-APR-2002.

PF 11-OCT-2001; 2001WO-US011606.

PR 12-OCT-2000; 2000US-00689123.

PA (UNITW) UNIV WASHINGTON.

PA (MEGA-) MEGAN HEALTH INC.

PA (CURT/) CURTIS R.

PI Curtiss R, Tinge SA;

DR WPI; 2002-444150/47.

XX Composition comprising microbe having attenuating mutation that comprises

PT insertion sequence containing recombinant transcription terminator,

PT useful as vaccine, and for delivering a desired gene product to  
PT individual.

XX Example 5; Fig 11; 91pp; English.

XX The present sequence is that of DNA located 5' to the phoA coding  
CC sequence of Escherichia coli. The sequence was used to illustrate an  
CC example of the invention relating to deletion of the phoA gene and  
CC replacement with a trpA terminator. The phoA mutation was introduced into  
CC Salmonella typhimurium SL1344 to produce strain MGN-1362. The present  
CC invention is based on the discovery that transcription terminators (TTs)  
CC can be inserted in a bacterial gene in an attenuating strategy which not  
CC only attenuates the virulence of the bacteria, but also serves to  
CC restrict the effect of the attenuating mutation to the target gene or  
CC operon of the bacteria. A claimed vaccine comprises a microorganism  
CC having an attenuating mutation in a chromosomal gene, the mutation  
CC comprising an insertion sequence which contains a recombinant TT, such as  
CC trpA. The microorganism is preferably Salmonella, Shigella or  
CC Escherichia, and the TT is preferably inserted in the phoA gene. The  
CC vaccine may further comprise a recombinant gene encoding a desired gene  
CC product from a virus, bacterium, protozoan, parasite or fungus, or  
CC encodes an autoantigen, gamete-specific antigen or an allergen. The  
CC attenuated microorganism is also used in a claimed method for delivering  
CC a desired gene product to an individual

XX Sequence 41 BP; 10 A; 13 C; 4 G; 14 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 28.8%; Score 40; DB 6; Length 41;

XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 11 TCGACGACTTAATATATGCTTCCTCACCCTTTTCTT 50

XX 1 TCGACGACTTAATATATGCTTCCTCACCCTTTTCTT 40

XX Db

XX RESULT 4  
XX ACA49482  
XX ID ACA49482 standard; DNA; 1385 BP.

XX AC ACA49482;

XX DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #31139.

XX KM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KM drug design; gene.

XX OS Salmonella paratyphi.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WC-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;

XX PI Wall D, Treweek JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX P-PSDB; ABU45612.

XX PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.  
XX Claim 14; SEQ ID NO 37352; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway of  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-regulated gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than S. aureus, S. typhimurium,  
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1385 BP; 353 A; 372 C; 335 G; 321 T; 0 U; 4 Other;

XX Query Match

XX Best Local Similarity 26.8%; Score 37.2; DB 8; Length 1385;

XX Matches 39; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX 1 GTGACTGTGTCGACGACTTAATATATGCTTCCTCACCCT 42

XX Db 1344 GTGACTGTGTCGACGACTTAATATATGCTTCCTCACCCT 1385

XX RESULT 5

XX ID ADE87477 standard; DNA; 266145 BP.

XX AC ADE87477;

XX DT 29-JAN-2004 (first entry)

XX DE Fowlpox virus genome DNA.

XX KM Fowlpox virus; FPV; virulence; tuberculostatic; protozoacide; antipylretic;

XX KM cytotactic; hepatotropic; antibacterial; vaccine; malaria; tuberculostatic;

XX KM East Coast fever; avipox virus; influenza; hepatitis;

XX KM human papilloma virus; tumour; leishmaniasis; listeriosis; theileria;

XX KM gene; ds.

XX OS Fowlpox virus.

XX PN WO2003047617-A2.

XX PD 12-JUN-2003.

XX PF 02-DEC-2002; 2002WC-GB005411.

XX PR 30-NOV-2001; 2001GB-00028733.

XX PR 30-NOV-2001; 2001US-0334649P.

XX	(ISIS-) ISIS INNOVATION LTD.
PA	
PI	Laidlaw S., Skinner M., Hill A., Gilbert S., Anderson R;
XX	
XX	WPI; 2003-513700/48.
DR	
XX	
PT	Treating and/or preventing e.g. malaria or tuberculosis, or eliciting an
PT	immune response, comprises administering a priming composition and a
PT	boosting composition containing a non-replicating viral vector in either
PT	order.
PS	Claim 30; SEQ ID NO 1; 302bp; English.
XX	
XX	The invention relates to a fowlpox virus (FPV) genome which has
CC	modifications in one or more wild-type FPV genes. The invention further
CC	relates to a novel method for treating and/or preventing a disease in a
CC	subject comprising administering two compositions, each containing a non-
CC	replicating viral vector. At least one of the compositions comprises a
CC	poxvirus vector derived from a fowlpox virus. The novel compositions have
CC	the following activities: vitruicide, tuberculostatic, protozoacide,
CC	antipruritic, cytostatic, hepatotropic, and antibacterial. The non-
CC	replicating viral vector is useful in a vaccine for an animal,
CC	particularly a mammal such as a primate, specifically human. The priming
CC	or boosting composition, or the kit is useful for manufacturing a
CC	medicament for treating and/or preventing a disease which is, or results
CC	from, a chronic infection such as malaria, tuberculosis or East Coast
CC	fever, or for eliciting a T-cell immune response in a subject. Non-
CC	cultured CEF cells are useful for growing an avipox virus, such as
CC	fowlpox virus. The method or the vaccine may further be used to treat or
CC	prevent influenza, hepatitis, human papilloma virus and other viral
CC	infections, malignancies such as tumours, leishmaniasis, listeriosis, and
CC	chellaria. This polynucleotide sequence represents the DNA of the fowlpox
CC	virus genome of the invention.
XX	
SQ	Sequence 266145 BP; 92375 A; 41026 C; 40846 G; 91897 T; 0 U; 1 Other;
	Query Match 23.5%; Score 32.6; DB 10; Length 266145;
	Best Local Similarity 55.9%; Pred. No. 8;
	Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
OY	18 ACTTAAATTAAGCGCTGCCTACCCCTTTTCCTTGAGAAGGCGTAGATTGTCTGGCT 77
DB	21181 AGTTACTTATGAGATATATTATTATCTTATTCGACGGAAGACAGAAAGTCATGGTATG 212404
OY	78 TTATTACTGTTTATCCCCAAGCACAATATCAAGCGTAGACTGTTCTTA 128
DB	21241 TTATGATGTTCTGTCCATAAAGGATATGATCATCTATTAATGTTTTTA 212931
RESULT 6	
ABQ72598/C	
ID	ABQ72598 standard; CDNA; 772 BP.
XX	
AC	ABQ72598;
XX	
DT	03-SEP-2002 (first entry)
XX	
DE	Human MDDT encoding CDNA SEQ ID NO 150.
XX	
KM	Human; MDDT; disease detection and treatment molecule polynucleotide;
KM	proliferative disorder; hepatitis; psoriasis; cancer; AIDS;
KM	autoimmune disorder; inflammatory disorder; allergy; multiple sclerosis;
KM	rheumatoid arthritis; transgenic; gene therapy; antiarteriosclerotic;
KM	hepatocarcinoma; antineoplastic; antipneumonia; cytostatic; anti-HIV;
KM	anti-allergic; antihaemic; antiaesthetic; antitachycardiac; anti-gout;
KM	neuroprotective; antirheumatic; antiarthritic; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200240715-A2.
XX	
PD	23-MAY-2002.

06-SEP-2001; 2001WO-US027628.

05-SEP-2000; 2000US-0229747P.  
05-SEP-2000; 2000US-0229748P.  
05-SEP-2000; 2000US-0229749P.  
05-SEP-2000; 2000US-0229750P.  
05-SEP-2000; 2000US-0229751P.  
05-SEP-2000; 2000US-0230583P.  
05-SEP-2000; 2000US-0230585P.  
05-SEP-2000; 2000US-0230514P.  
05-SEP-2000; 2000US-0230515P.  
05-SEP-2000; 2000US-0230517P.  
05-SEP-2000; 2000US-0230518P.  
05-SEP-2000; 2000US-0230519P.  
05-SEP-2000; 2000US-0230597P.  
05-SEP-2000; 2000US-0230598P.  
05-SEP-2000; 2000US-0230599P.  
05-SEP-2000; 2000US-0230610P.  
05-SEP-2000; 2000US-0230865P.  
05-SEP-2000; 2000US-0230988P.  
05-SEP-2000; 2000US-0230989P.  
07-SEP-2000; 2000US-0230951P.  
07-SEP-2000; 2000US-0231163P.  
07-SEP-2000; 2000US-0231167P.

(INCYT) INCYTE GENOMICS INC.

William S, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL, Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Dahl CK, Momiya MG, Bradley DL, Rohatgi SD, Harris B, Rosdebery AM, Gerstin EH, Petalca CH, David MH, Panzer SR, Daffo A, Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR; MPI; 2002-527544/56.  
P-PSDB; ABP51381.

Novel human disease detection and treatment polypeptide, useful in diagnosis, prevention or treatment of cell proliferative disorders e.g. arteriosclerosis, cirrhosis and an autoimmune/inflammatory disorder e.g. AIDS.

Claim 1; Page 390-391; 618pp; English.

The invention relates to an isolated human disease detection and treatment (MDT) polypeptide (I) selected from a polypeptide having a sequence selected from 254 sequences (ABP51231-ABP51484) given in the specification, a naturally occurring polypeptide comprising a sequence having at least 90% identity to (I) or a biologically active or immunogenic fragment of (I). (I) is useful for screening a compound for effectiveness as an agonist or antagonist, for screening a compound that specifically binds (I) or modulates the activity of (I), and for preparing a polyclonal or monoclonal antibody by hybridoma technology. Nucleic acids (II) (ABQ72449-ABQ72700) encoding (I) are useful for screening a compound for effectiveness in altering expression of a target polynucleotide comprising. Oligonucleotides and antibodies are useful for detecting MDT in a sample or for assessing toxicity of a test compound, in a diagnostic test for a condition or a disease associated with the expression of MDT in a biological sample, for detecting (I) in a sample, and for purifying (I) from a sample. A composition comprising (I), an agonist or antagonist is useful for treating a disease or condition associated with decreased or increased expression of functional MDT. (I) or (II) are useful for diagnosing, treating or preventing disorders associated with aberrant expression of MDT, where the disorders are selected from a cell proliferative disorder such as arteriosclerosis, cirrhosis, hepatitis, psoriasis, and cancer and an autoimmune/inflammatory disorder such as AIDS, Addison's disease, allergy, anemias, asthma, atherosclerosis, gout, multiple sclerosis or rheumatoid arthritis. (II) are useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germ-line gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to

CC translocation or inversion among normal, carrier or affected individuals  
CC and as hybridisation probes for mapping naturally occurring genomic  
CC sequences

XX Sequence 772 BP; 252 A; 147 C; 156 G; 217 T; 0 U; 0 Other;

XX Query Match 22.3%; Score 31; DB 6; Length 772;

XX Best Local Similarity 57.9%; Pred. No. 5.1;

XX Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

XX Db 45 TTCTCTCAGAAAGAGGCTGCTGTTTATTTATCTGTTATCCCAAGCACC 104

XX 549 TTCCATATAAAAAAATCCCTCCATTAGTTATTATATACAGTTTGTCCAAAGCGCT 490

XX 105 ATATACACGCTGACTGTTCTTATTTATTAACA 139

XX 489 TTCTGTAATGCTGAAATCTCTCAGTCAACACA 455

XX Db

XX RESULT 7

XX ABQ72685/c

XX ID ABQ72685 standard; cDNA; 810 BP.

XX AC ABQ72685;

XX DT 03-SEP-2002 (first entry)

XX DE Human MDDT encoding cDNA SEQ ID NO 237.

XX XX Human: MDDT; disease detection and treatment molecule polynucleotide;

XX XX proliferative disorder; hepatitis; psoriasis; cancer; AIDS;

XX XX rheumatoid arthritis; transgenic; gene therapy; antileukosclerotic;

XX XX hepatocellular carcinoma; antiproliferative; cytostatic; anti-HIV;

XX XX antiangiogenic; antineoplastic; antitumor; antileukosclerotic; antitumor;

XX XX neuroprotective; antineoplastic; antileukosclerotic; gene; ss.

XX OS Homo sapiens.

XX PN WO200240715-A2.

XX PD 23-MAY-2002.

XX PF 06-SEP-2001; 2001WO-US027628.

XX PR 05-SEP-2000; 2000US-0229747P.

XX PR 05-SEP-2000; 2000US-0229748P.

XX PR 05-SEP-2000; 2000US-0229750P.

XX PR 05-SEP-2000; 2000US-0229751P.

XX PR 05-SEP-2000; 2000US-0230583P.

XX PR 06-SEP-2000; 2000US-0230585P.

XX PR 06-SEP-2000; 2000US-0230587P.

XX PR 06-SEP-2000; 2000US-0230589P.

XX PR 06-SEP-2000; 2000US-0230591P.

XX PR 06-SEP-2000; 2000US-0230593P.

XX PR 06-SEP-2000; 2000US-0230595P.

XX PR 06-SEP-2000; 2000US-0230597P.

PI Dahl CR, Momiyama NG, Bradley DL, Rohatgi SD, Harris B,  
PI Rosberry AM, Gershtin EH, Peralta CH, David MH, Panzer SR, Flores V,  
PI Daffo A, Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;  
PI WPI; 2002-527544/56.  
PI P-PSDB; ABP51469.

XX Novel human disease detection and treatment polypeptide, useful in  
XX diagnosis, prevention or treatment of cell proliferative disorders e.g.  
XX arteriosclerosis, cirrhosis and an autoimmune/inflammatory disorder e.g.  
XX AIDS.

XX Claim 1, Page 435, 618pp; English.

XX The invention relates to an isolated human disease detection and  
XX treatment (MDDT) polypeptide (I) selected from a polypeptide having a  
XX sequence selected from 254 sequences (ABP51231-ABP51484) given in the  
XX specification, a naturally occurring polypeptide comprising a sequence  
XX having at least 90% identity to (I) or a biologically active or  
XX immunogenic fragment of (I). (I) is useful for screening a compound for  
XX effectiveness as an agonist or antagonist, for screening a compound that  
XX specifically binds (I) or modulates the activity of (I), and for  
XX preparing a polyclonal or monoclonal antibody by hybridoma technology.

XX Nucleic acids (II) (ABQ72449-ABQ72700) encoding (I) are useful for  
XX screening a compound for effectiveness in altering expression of a target  
XX polynucleotide comprising. Oligonucleotides and antibodies are useful for  
XX detecting MDDT in a sample or for assessing toxicity of a test compound,  
XX in a diagnostic test for a condition or a disease associated with the  
XX expression of MDDT in a biological sample, for detecting (I) in a sample,  
XX and for purifying (I) from a sample. A composition comprising (I), an  
XX agonist or antagonist is useful for treating a disease or condition  
XX associated with decreased or increased expression of functional MDDT. (I)  
XX or (II) are useful for diagnosing, treating or preventing disorders  
XX associated with aberrant expression of MDDT, where the disorders are  
XX selected from a cell proliferative disorder such as arteriosclerosis,  
XX cirrhosis, hepatitis, psoriasis, and cancer and an  
XX autoimmune/inflammatory disorder such as AIDS, Addison's disease,  
XX allergy, anaemia, asthma, atherosclerosis, gout, multiple sclerosis or  
XX rheumatoid arthritis. (II) are useful for creating knockin humanised  
XX animals or transgenic animals to model human diseases, in somatic or  
XX germine gene therapy, to generate a transcript image of a tissue or cell  
XX type, for detecting differences in the chromosomal location due to  
XX translocation or inversion among normal, carrier or affected individuals  
XX and as hybridisation probes for mapping naturally occurring genomic  
XX sequences

XX Sequence 810 BP; 262 A; 153 C; 165 G; 230 T; 0 U; 0 Other;

XX Query Match 22.3%; Score 31; DB 6; Length 810;

XX Best Local Similarity 57.9%; Pred. No. 5.2;

XX Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

XX Db 45 TTCTCTCAGAAAGAGGCTGCTGTTTATTTATCTGTTATCCCAAGCACC 104

XX 549 TTCCATATAAAAAAATCCCTCCATTAGTTATTATATACAGTTTGTCCAAAGCGCT 490

XX 105 ATATACACGCTGACTGTTCTTATTTATTAACA 139

XX 489 TTCTGTAATGCTGAAATCTCTCAGTCAACACA 455

XX Db

XX RESULT 8

XX AAK81193/c

XX ID AAK81193 standard; DNA; 19521 BP.

XX AC AAK81193;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36005.

XX XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX KM

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Homo sapiens.  
 OS  
 XX WO200157182-A2.  
 PN  
 XX  
 XX  
 PD 09-AUG-2001.  
 PF 17-JAN-2001; 2001WO-US001354.  
 XX  
 PR 31-JAN-2000; 2000US-0179665P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-020515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-022547P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226688P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 06-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235848P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237033P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246522P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.

XX



PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI, 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
PS Disclosure; SEQ ID NO 36005; 3071bp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK7694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 19521 BP; 5672 A; 3465 C; 3902 G; 6482 T; 0 U; 0 Other;  
XX  
Query Match 22.2%; Score 30.8; DB 4; Length 19521;  
Best Local Similarity 63.5%; Pred. No. 15;  
Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
XX  
QY 1 GTGACTCTGCTGACGCACTTAATAATGCGCTCCTCCTTTCTTCAGAAAGAG 60  
DB 1311 GTAACTCTGGGCAACTTACATAAATCTCTGTCCTCAGCTTCGACCTGAGAAAGG 1252  
XX  
QY 61 GTGACTATTGTCT 74  
DB 1251 CTAACTTCTTAT 1238  
XX  
RESULT 9  
ADA41637/c  
ID ADA41637 standard; DNA; 19521 BP.  
XX  
AC ADA41637;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human secreted protein related DNA.  
XX  
KW Human; secreted protein; cancer; hyperproliferative disorder;  
KW rheumatoid arthritis; autoimmune disorder; hematopoietic disorder;  
KW anaemia; allergic reaction; asthma; cardiovascular disorder;  
KW wound healing; cytostatic; immunosuppressive; nootropic; neuroprotective;  
KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;  
KW vulnary; cardiac; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2002102993-A2.  
XX  
PD 27-DEC-2002.  
XX  
PF 19-MAR-2002; 2002WO-US008123.  
XX  
PR 21-MAR-2001; 2001US-0277340P.  
PR 19-JUL-2001; 2001US-0306171P.  
PR 13-NOV-2001; 2001US-0331287P.  
XX

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI, 2003-175238/17.  
XX  
PT New human secreted proteins and nucleic acid molecules, useful for  
XX preparing a diagnostic or pharmaceutical composition for diagnosing,  
XX preventing or treating cancer or other hyperproliferative disorder,  
XX asthma, allergies or AIDS.  
XX  
PS Disclosure; SEQ ID NO 2020; 3205bp; English.  
XX  
CC The invention relates to novel genes ADA39629-ADA40565 and proteins  
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,  
CC treating or ameliorating medical conditions e.g. by protein or gene  
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their  
CC fragments, and agonists or antagonists that bind to the polypeptide are  
CC useful for preparing a diagnostic or pharmaceutical composition for  
CC diagnosing or treating cancer or other hyperproliferative disorder. The  
CC polypeptides and nucleic acid molecules are also useful for detecting,  
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer  
CC or other hyperproliferative disorders including neoplasms, autoimmune  
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus  
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic  
CC anaemia), hematopoietic or hematological disorders (e.g. anaemia,  
CC thrombocytopenia), allergic reactions including asthma or eczema,  
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory  
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.  
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders  
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,  
CC fungal or viral infections including HIV/AIDS), or wound healing and  
CC disorders of epithelial cell proliferation. The nucleic acids are also  
CC useful for chromosome identification, radiation hybrid mapping or long-  
CC range restriction mapping, as molecular weight markers, or as  
CC hybridization or diagnostic probes. The polypeptides and antibodies are  
CC useful for providing immunological probes for differential identification  
CC of the tissues immunohistochemistry assays. The sequences given in  
CC ADA40803-ADA41665 represent DNA sequences related to human secreted  
CC proteins. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX  
SQ Sequence 19521 BP; 5672 A; 3465 C; 3902 G; 6482 T; 0 U; 0 Other;  
XX  
Query Match 22.2%; Score 30.8; DB 8; Length 19521;  
Best Local Similarity 63.5%; Pred. No. 15;  
Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
XX  
QY 1 GTGACTCTGCTGACGCACTTAATAATGCGCTCCTCCTTTCTTCAGAAAGAG 60  
DB 1311 GTAACTCTGGGCAACTTACATAAATCTCTGTCCTCAGCTTCGACCTGAGAAAGG 1252  
XX  
QY 61 GTGACTATTGTCT 74  
DB 1251 CTAACTTCTTAT 1238  
XX  
RESULT 10  
ADA57769/c  
ID ADA57769 standard; DNA; 19521 BP.  
XX  
AC ADA57769;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE BAC fragment containing human secreted protein gene #505.  
XX  
KW Immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;  
KW cytotoxic; cerebroprotective; neuroprotective; nootropic;  
KW cardiovascular; antiarteriosclerotic; gene therapy;  
KW human secreted protein; immune disorder; inflammation;  
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;

KM inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
 KM multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
 KM Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
 KM triple helix formation; antisense gene therapy; forensic biology; ds;  
 KM gene; bacterial artificial chromosome.  
 OS Homo sapiens.  
 XX WO2002102994-A2.  
 XX PN 27-DEC-2002.  
 XX PD 19-MAR-2002; 2002WO-US008278.  
 XX PF 21-MAR-2001; 2001US-0277340P.  
 XX PR 19-JUL-2001; 2001US-0306172P.  
 XX PR 13-NOV-2001; 2001US-0331287P.  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Ruben SM;  
 XX DR WPI; 2003-167512/16.  
 XX PT New human secreted polypeptides and polynucleotides, useful for  
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory  
 PT conditions, respiratory disorders, cancers, CNS disorders, or  
 PT neurodegenerative disorders.  
 XX PS Disclosure; SEQ ID NO 1962; 1754pp; English.  
 XX XX The invention relates to 592 new human secreted polypeptides useful for  
 CC diagnosing, treating or preventing e.g. immune disorders, inflammatory  
 CC conditions, respiratory disorders, cancers, CNS disorders, or  
 CC neurodegenerative disorders, or polypeptides comprising an amino acid  
 CC sequence at least 95% identical to the new sequences. The polypeptides,  
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic  
 CC acids encoding the polypeptides, agonists or antagonists that binds to  
 CC the polypeptides, are useful in preparing diagnostic or pharmaceutical  
 CC compositions for diagnosing, treating or preventing an e.g. immune  
 CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,  
 CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and  
 CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders  
 CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative  
 CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and  
 CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The  
 CC polynucleotides are useful for chromosome identification, chromosome  
 CC mapping, for controlling gene expression through triple helix formation  
 CC or antisense DNA or RNA, in gene therapy, for identifying individuals  
 CC from minute biological samples, in forensic biology, and as hybridization  
 CC probes. The polypeptides are useful for as molecular weight markers on  
 CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)  
 CC gels, to raise antibodies, for testing biological activities, and for  
 CC treating or preventing neural disorders, immune system disorders,  
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds  
 CC to a bacterial artificial chromosome (BAC) fragment containing the gene  
 CC encoding one of the polypeptides of the invention. Note: The sequence  
 CC data for this patent did form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp://ipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 19521 BP; 5672 A; 3465 C; 3902 G; 6482 T; 0 U; 0 Other;  
 XX  
 XX Query Match 22.2%; Score 30.8; DB 10; Length 19521;  
 XX Best Local Similarity 63.5%; Pred. No. 15;  
 XX Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
 QY 1 GTGACTTCGCTGACGACAACTTAATAATGCCGCTCACCCTTTCTTCGAGAAAGG 60  
 XX |||||||  
 DB 1311 GTAACTCTGGGCAACTTACATAAATCTGTGCTCAGCTTGCACCTGAGAAAAGG 1252  
 XX |||||||  
 QY 61 GTGACTATTGTCT 74

DB 1251 CTAACAATTCCTAT 1238  
 |||||  
 RESULT 11  
 ID ACA02032 standard; DNA; 1539 BP.  
 XX ACA02032;  
 AC 04-JUN-2003 (first entry)  
 XX DT  
 XX C. glutamicum derived ORF SEQ ID 2023.  
 DE  
 XX Corynebacterium derived ORF SEQ ID 2023.  
 XX Coryneform; nucleic acid array; fermentation; culture; ds.  
 KM Corynebacterium glutamicum.  
 XX OS Corynebacterium glutamicum.  
 XX DE10128510-A1.  
 XX PN 19-DEC-2002.  
 XX PD 13-JUN-2001; 2001DE-01028510.  
 XX PF 13-JUN-2001; 2001DE-01028510.  
 XX PR 13-JUN-2001; 2001DE-01028510.  
 XX PA (DEGS) DEGUSA AG.  
 XX PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;  
 XX DR WPI; 2003-279970/28.  
 XX PT New nucleic acid array useful for monitoring mRNA expression of  
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid  
 PT from Corynebacterium glutamicum.  
 XX PS Claim 1; Page 665-666; 709pp; German.  
 XX XX This invention describes a novel nucleic acid array involving  
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to  
 CC analyze C. glutamicum, particularly for monitoring a fermentation process  
 CC to determine expression levels of C. glutamicum cellular mRNA. Such  
 CC monitoring particularly differentiates between expression levels of  
 CC different strains of C. glutamicum and allows the adjustment of different  
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.  
 CC glutamicum derived polynucleotides described in the disclosure of the  
 CC invention  
 XX SQ Sequence 1539 BP; 393 A; 394 C; 414 G; 338 T; 0 U; 0 Other;  
 XX  
 XX Query Match 22.0%; Score 30.6; DB 8; Length 1539;  
 XX Best Local Similarity 53.8%; Pred. No. 8.4;  
 XX Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
 QY 17 AACTTAATTAATGCTGCTCACTCTTTCTTCGAGAAAGGCTACTATTTCTCGG 76  
 XX |||||||  
 DB 241 AACGCTATTGATGGCAGGAGGAGCATGCTTATGAGAGAACTCAACCCATTGCTGGG 300  
 XX |||||||  
 QY 77 TTATTAATCTGTTATCCCAAGCAGCATATATCAAGCCTAGACTGTTCTATTGTT 133  
 XX |||||||  
 DB 301 ATCATGACAGATGTCTGTTCAACACCAATTCACCATCCACGATATATCGGT 357  
 XX |||||||  
 RESULT 12  
 AAH67957  
 ID AAH67957 standard; DNA; 2799 BP.  
 XX AAH67957;  
 AC 26-SEP-2001 (first entry)  
 XX DT  
 XX C glutamicum coding sequence fragment SEQ ID NO: 2992.  
 XX DE

KM	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP1108790-A2.
XX	
PD	20-JUN-2001.
XX	
PF	18-DEC-2000; 2000EP-00127688.
XX	
PR	16-DEC-1999; 99JP-00377484.
PR	07-APR-2000; 2000JP-00159162.
PR	03-AUG-2000; 2000JP-00280988.
XX	
PA	(KYOW ) KYOWA HAKKO KOGYO KK.
XX	
PI	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX	
DR	WPI; 2001-376931/40.
XX	P-PSDB; AAG92738.
XX	
PT	Novel polynucleotides derived from Coryneform bacteria, for identifying
PT	mutation point of a gene, measuring expression of a gene, analyzing
PT	expression profile or pattern of a gene and identifying homologous gene.
XX	
PS	Claim 8; SEQ ID NO 2992; 246bp + Sequence Listing; English.
XX	
CC	The present invention provides a number of nucleotide and protein
CC	sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC	are useful for identifying the mutation point of a gene derived from a
CC	mutant of coryneform bacterium, measuring expression amount and analysing
CC	the expression profile or expression pattern of a gene derived from
CC	Coryneform bacterium, and identifying a homologue of a gene derived from
CC	Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC	acids, nucleic acids, vitamins, saccharides and organic acids,
CC	particularly L-lysine. The present sequence is a nucleic acid described
CC	in the exemplification of the invention. Note: The sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from the European Patent Office
XX	
SQ	Sequence 2799 BP; 715 A; 721 C; 728 G; 625 T; 0 U; 0 Other;
XX	
Query Match	22.0%; Score 30.6; DB 5; Length 2799;
Best Local Similarity	53.8%; Pred. No. 10;
Matches 63; Conservative	0; Mismatches 54; Indels 0; Gaps 0;
OY	17 AACTTAATAATGACCTGCCTCACCTCTTTCTTCAGAAAGAGGCTGACTATTGTCTGCG 76
DB	211 AACGCATGATGATGAGCAGGAGGATGCTTATGAGAAAGAACTAACCCATGTTGGG 330
OY	77 TTTATTAAGTGTTAATCCCAAGACCAATTAACGCTGATGTTCTTATTGTT 133
DB	331 ATCATGACAGAGTGCTGTTTCAACACCAAAATTCAACGATCAACAGTATGATCGST 387
RESULT 13	
ID	AAH68533/C
XX	AAH68533 standard; DNA; 349980 BP.
AC	AAH68533;
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum coding sequence fragment SEQ ID NO: 7068.
XX	
KW	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP1108790-A2

XX 20-JUN-2001.  
ED XX  
PF 18-DEC-2000; 2000EP-00127688.  
XX  
PR 16-DEC-1999; 99JP-00377484.  
XX  
PR 07-APR-2000; 2000JP-00159162.  
PR 03-AUG-2000; 2000JP-00280988.  
XX  
PA (KYOW ) KYOMA HAKKO KOGYO KK.  
XX  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A,  
XX  
DR WPI; 2001-376931/40.  
XX  
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX  
PS Disclosure; SEQ ID NO 7068; 246bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of nucleotide and protein  
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of corynebacterium bacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Corynebacterium bacterium, and identifying a homologue of a gene derived from  
CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;  
XX  
Query Match 22.0%; Score 30.6; DB 5; Length 349980;  
Best Local Similarity 53.8%; Pred. No. 39;  
Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
QY 17 AACTTAATAATATGCTGCTCAACCTCTTTTCTTCAGAAAGGGTGAATTTGCTG 76  
DB 200056 AACGTCATTGATGAGCGAAGGCGATGCTTCATGCAAGAACTCAACCCATTGGTGG 199997  
QY 77 TTATTAAGTGTATATCCCAAGACCATATCAAGCTAGACTGTTCTTATGTT 133  
DB 199996 ATCATGACAGGTGTCTGTTTCACACCAATTCACCGATCAGCTATGATGCT 199940  
RESULT 14  
ID AAD63515/C  
XX AAD63515 standard; DNA; 31241 BP.  
XX  
AC AAD63515;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Mycoplasma genitalium gene involved in transport and binding.  
XX  
XX Genetic operating system; nanomachine genome; bioresactor; bioremediation;  
KW therapeutic biomolecule; energy conversion system; processing system;  
KW anabolic; catabolic system; biological film; cosmetic application;  
KW coating; de.  
OS Mycoplasma genitalium.  
XX  
XX US2003138777-A1.  
XX  
XX 24-JUL-2003.  
XX  
XX 20-SEP-2001; 2001US-00960858.  
XX  
XX PF

```

PR 20-SEP-2001; 2001US-00960858.
XX
XX (EVAN/) EVANS G A.
XX
XX Evans GA;
XX
XX WPI; 2003-851721/79.
XX
XX Basic genetic operating system for an autonomous prototrophic
XX nanomachine, comprises a nanomachine genome encoding a minimal gene set.
XX
XX Example 1; Page 127-141; 170pp; English.
XX
XX The invention relates to a basic genetic operating system which comprises
XX a nanomachine genome encoding a minimal gene set for viability. The basic
XX genetic operating system is used for an autonomous prototrophic
XX nanomachine or autonomous auxotrophic nanomachine. The nanomachine is
XX used as a bioreactor, for bioremediation, for production of a therapeutic
XX biomolecule or as a diagnostic reagent, for production of a diagnostic
XX indicator or as a diagnostic reagent, as a delivery system, as an
XX artificial tissue or organ system, an energy conversion system, as a
XX processing system, as an anabolic or catabolic system, for production of
XX biological films or coatings that may respond to the environment and for
XX cosmetic applications including pharmaceuticals. The present sequence is
XX Mycoplasma genitalium nanomachine gene
SQ Sequence 31241 BP; 10733 A; 4570 C; 5361 G; 10577 T; 0 U; 0 Other;

Query Match 21.9%; Score 30.4; DB 10; Length 31241;
Best Local Similarity 55.8%; Pred. No. 23;
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 34 CCTCACCCCTTTCTTCAGAAAGGGTGACTATTGTCGTGTTATTACTGTTATC 93
DB 25866 CCTCAGCCAAATTTTATTACCAAGTGGAGATTTCACATCGTTTGTCTTGTGC 25807
QY 94 CCCAAGCACCATAATCAAGCTAGACTGTTCTTATTGTTACA 137
DB 25806 ATTAAGAGAACGCTATCAATCTATACCCCTTTTGTCTAATA 25763

RESULT 15
ACC69145/C
ID ACC69145 standard; DNA; 31241 BP.
XX
XX ACC69145;
XX
XX 10-JUL-2003 (first entry)
XX
XX M. genitalium transport and binding gene cassette DNA SEQ ID NO:13.
XX
XX Mycoplasma genitalium; gene cassette; replication; transcription;
XX translation; metabolism; basic genetic operating system; gene therapy;
XX autonomous prototrophic nanomachine; auxotrophic nanomachine;
XX nanomachine; bioreactor; bioremediation; therapeutic; delivery system;
XX artificial tissue; artificial organ system; energy conversion system;
XX processing system; anabolic system; catabolic system; biological film;
XX biological coating; cosmetic; gene; ds.
XX
XX Mycoplasma genitalium.
XX
XX WO2003025145-A2.
XX
XX 27-MAR-2003.
XX
XX 18-SEP-2002; 2002WO-US029811.
XX
XX 20-SEP-2001; 2001US-00960870.
XX
XX (EGBA-) EGBA BIOSCIENCES INC.
XX
XX Evans GA;
XX
XX

```

```

DR WPI; 2003-354602/33.
XX
XX New basic genetic operating system for autonomous prototrophic or
XX auxotrophic nanomachine, useful for therapeutic, diagnostic or industrial
XX purposes, comprises a nanomachine genome encoding a gene set for
XX viability or replication.
XX
XX Example 1; Page 231-240; 250pp; English.
XX
XX The present invention describes a basic genetic operating system for an
XX autonomous prototrophic or auxotrophic nanomachine comprising a
XX nanomachine genome encoding a minimal gene set sufficient for viability
XX or replication, optionally in the presence of an auxotrophic molecule.
XX Also described is an autonomous prototrophic or auxotrophic nanomachine
XX comprising a basic genetic operating system for autonomous prototrophic
XX or auxotrophic viability or replication, optionally in the presence of an
XX auxotrophic molecule, and a particle envelope. The nanomachines can be
XX used in gene therapy. The basic genetic operating system or nanomachine
XX is useful in therapeutic, diagnostic and industrial applications, e.g. as
XX a bioreactor, for bioremediation, for the production of a therapeutic
XX biomolecule or as a therapeutic reagent, for the production of a
XX diagnostic indicator or reagent, as a delivery system, as an artificial
XX tissue or organ system, as an energy conversion system, as a processing
XX system, as an anabolic or catabolic system, for the production of
XX biological films or coatings, and for cosmetic applications. The present
XX sequence represents a Mycoplasma genitalium gene cassette nucleotide
XX sequence, which is used in an example from the present invention for the
XX design and synthesis of a basic genetic operation system for a
XX replication competent nanomachine
SQ Sequence 31241 BP; 10733 A; 4570 C; 5361 G; 10577 T; 0 U; 0 Other;

Query Match 21.9%; Score 30.4; DB 10; Length 31241;
Best Local Similarity 55.8%; Pred. No. 23;
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 34 CCTCACCCCTTTCTTCAGAAAGGGTGACTATTGTCGTGTTATTACTGTTATC 93
DB 25866 CCTCAGCCAAATTTTATTACCAAGTGGAGATTTCACATCGTTTGTCTTGTGC 25807
QY 94 CCCAAGCACCATAATCAAGCTAGACTGTTCTTATTGTTACA 137
DB 25806 ATTAAGAGAACGCTATCAATCTATACCCCTTTTGTCTAATA 25763

Search completed: August 3, 2005, 21:57:32
JOD time : 282 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:54:05 ; Search time 358 Seconds  
(without alignments)  
2513.010 Million cell updates/sec

Title: US-10-070-882a-2

Perfect score: 139  
Sequence: 1 gtgactctgcgtcagcaact.....ctgtctctatgttaacaca 139

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.2	26.8	1385	17	US-10-282-122A-37352
2	33	23.7	100596	22	US-10-737-082-92
3	33	23.7	100596	22	US-10-765-790-92
4	32.6	23.5	266145	21	US-10-856-118-1
5	31	22.3	772	19	US-10-363-829-150
6	31	22.3	810	19	US-10-363-829-237
7	31	22.3	2337	21	US-10-486-706-223

8	30.6	22.0	2799	9	US-09-738-626-2992	Sequence 2992, App
9	30.6	22.0	3309400	9	US-09-738-626-1	Sequence 1, Appl
10	30.4	21.9	15649	22	US-10-737-082-4	Sequence 4, Appl
11	30.4	21.9	15649	22	US-10-765-790-4	Sequence 4, Appl
12	30.4	21.9	31241	10	US-09-960-870-13	Sequence 13, Appl
13	30.4	21.9	31241	10	US-09-960-858-13	Sequence 13, Appl
14	30.4	21.9	31241	18	US-10-251-668-13	Sequence 13, Appl
15	30.4	21.9	104644	19	US-10-433-287-79	Sequence 79, Appl
16	30.4	21.9	109559	19	US-10-322-281-137	Sequence 137, App
17	30.2	21.7	130349	21	US-10-741-600-17619	Sequence 17619, App
18	30.2	21.7	138363	19	US-10-367-094-117	Sequence 117, App
19	30.2	21.6	458	18	US-10-424-559-105245	Sequence 105245, App
20	30	21.6	622	19	US-10-600-070-201	Sequence 201, App
21	29.8	21.4	301	9	US-09-759-143-293	Sequence 293, App
22	29.8	21.4	301	9	US-09-780-665-293	Sequence 293, App
23	29.8	21.4	301	9	US-09-822-827-293	Sequence 293, App
24	29.8	21.4	301	9	US-09-232-880-293	Sequence 293, App
25	29.8	21.4	301	9	US-09-895-793-293	Sequence 293, App
26	29.8	21.4	301	9	US-09-895-814-293	Sequence 293, App
27	29.8	21.4	301	13	US-10-012-896-293	Sequence 293, App
28	29.8	21.4	301	14	US-10-010-940-293	Sequence 293, App
29	29.8	21.4	301	16	US-10-144-678A-293	Sequence 293, App
30	29.8	21.4	301	16	US-10-294-025-293	Sequence 293, App
31	29.8	21.4	340	20	US-10-357-930-7475	Sequence 7475, App
32	29.8	21.4	409	20	US-10-357-930-37436	Sequence 37436, App
33	29.8	21.4	434	15	US-10-102-524-107	Sequence 307, App
34	29.8	21.4	434	15	US-10-102-524-540	Sequence 540, App
35	29.8	21.4	444	13	US-10-027-632-81146	Sequence 81146, App
36	29.8	21.4	444	17	US-10-027-632-81146	Sequence 81146, App
37	29.8	21.4	463	13	US-10-027-632-109769	Sequence 109769, App
38	29.8	21.4	463	17	US-10-027-632-109769	Sequence 109769, App
39	29.8	21.4	492	9	US-09-954-456-46	Sequence 46, Appl
40	29.8	21.4	492	9	US-09-954-456-962	Sequence 962, App
41	29.8	21.4	492	21	US-10-843-641A-3073	Sequence 3073, App
42	29.8	21.4	492	21	US-10-843-641A-3989	Sequence 3989, App
43	29.8	21.4	617	13	US-10-027-632-229999	Sequence 229999, App
44	29.8	21.4	617	17	US-10-027-632-229999	Sequence 229999, App
45	29.8	21.4	1400	21	US-10-956-157-7378	Sequence 7378, App

#### ALIGNMENTS

RESULT 1  
US-10-282-122A-37352  
; Sequence 37352, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Lianggu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Foreyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 37352  
LENGTH: 1385  
TYPE: DNA  
ORGANISM: *Salmonella paratyphi A*  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (72)..(72)  
OTHER INFORMATION: n=g, a, t or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (74)..(75)  
OTHER INFORMATION: n=g, a, t or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (158)..(158)  
OTHER INFORMATION: n=g, a, t or c  
US-10-282-122A-37352

Query Match 26.8%; Score 37.2; DB 17; Length 1385;  
Best Local Similarity 92.9%; Pred. No. 0.074;  
Matches 39; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGACTCTGGTGAAGAACTTAATATGCGCTGACCCCT 42  
DB 1344 GGGACTCTGGTGAAGAACTTAATATGCGCTGACCCCT 1385

RESULT 2  
US-10-737-082-92/c  
Sequence 92, Application US/10737082  
Publication No. US20050130170A1  
GENERAL INFORMATION:  
APPLICANT: Bayer Healthcare LLC  
APPLICANT: Beard, Chris  
APPLICANT: Gannon, Allison  
APPLICANT: Harvey, Jeanne  
APPLICANT: Lechner, John F.  
APPLICANT: Li, Zheng  
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
FILE REFERENCE: 1657/2032  
CURRENT APPLICATION NUMBER: US/10/737,082  
CURRENT FILING DATE: 2003-12-16  
PRIOR APPLICATION NUMBER: US 10/737,082  
PRIOR FILING DATE: 2003-12-16  
NUMBER OF SEQ ID NOS: 300  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 92  
LENGTH: 100596  
TYPE: DNA  
ORGANISM: *Homo sapiens*  
US-10-737-082-92

Query Match 23.7%; Score 33; DB 22; Length 100596;  
Best Local Similarity 65.8%; Pred. No. 12;  
Matches 48; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 18 ACTTAATATGCGCTGACCCCTTTCTTCTTCAAGAGGGTACTATTGCTGGT 77

DB 34052 ATTGAATATGCTGTTTTCCTTCTTTTAAACAATAATGCTAATTAATGACTATT 33993  
QY 78 TTATTAAGTGT 90  
DB 33992 TTATTAAGTGT 33980

RESULT 3  
US-10-765-790-92/c  
Sequence 92, Application US/10765790  
Publication No. US20050130172A1  
GENERAL INFORMATION:  
APPLICANT: Bayer Healthcare LLC  
APPLICANT: Beard, Chris  
APPLICANT: Gannon, Allison  
APPLICANT: Harvey, Jeanne  
APPLICANT: Lechner, John F.  
APPLICANT: Li, Zheng  
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
FILE REFERENCE: 1657/2035  
CURRENT APPLICATION NUMBER: US/10/765,790  
CURRENT FILING DATE: 2004-01-27  
PRIOR APPLICATION NUMBER: US 10/737,082  
PRIOR FILING DATE: 2003-12-16  
NUMBER OF SEQ ID NOS: 300  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 92  
LENGTH: 100596  
TYPE: DNA  
ORGANISM: *Homo sapiens*  
US-10-765-790-92

Query Match 23.7%; Score 33; DB 22; Length 100596;  
Best Local Similarity 65.8%; Pred. No. 12;  
Matches 48; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 18 ACTTAATATGCGCTGACCCCTTTCTTCTTCAAGAGGGTACTATTGCTGGT 77  
DB 34052 ATTGAATATGCTGTTTTCCTTCTTTTAAACAATAATGCTAATTAATGACTATT 33993

QY 78 TTATTAAGTGT 90  
DB 33992 TTATTAAGTGT 33980

RESULT 4  
US-10-856-118-1  
Sequence 1, Application US/10856118  
Publication No. US20050025747A1  
GENERAL INFORMATION:  
APPLICANT: Laidlaw, Stephen  
APPLICANT: Skinner, Mike  
APPLICANT: Hill, Adrian V.S.  
APPLICANT: Gilbert, Sarah C.  
APPLICANT: Anderson, Richard  
TITLE OF INVENTION: Vaccine  
FILE REFERENCE: 3742.1000-000  
CURRENT APPLICATION NUMBER: US/10/856,118  
CURRENT FILING DATE: 2004-05-27  
PRIOR APPLICATION NUMBER: PCT/GB02/005411  
PRIOR FILING DATE: 2002-12-02  
PRIOR APPLICATION NUMBER: GB0128733.3  
PRIOR FILING DATE: 2001-11-30  
PRIOR APPLICATION NUMBER: US 60/334,649  
PRIOR FILING DATE: 2001-11-30  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 266145  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Fowlpox virus genome  
US-10-856-118-1

Query Match 23.5%; Score 32.6; DB 21; Length 266145;  
Best Local Similarity 55.9%; Pred. No. 24;  
Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 18 ACTTAATGAATGCGCTGACCCCTTTCTTGGAAGAGGGGAGCTATTGTCGTGT 77  
DB 21181 AGTACTATGATATATTTATATCTTATGACGAGGAGACGAAGATCCATGATGTG 21240

QY 78 TTATTAAGCTTTATCCCAAGACACATTAATCAAGCTAGCTGTTCTTA 128  
DB 21241 TTATGATGTTCTGTCATTAAGGATATGATCATATATATGTTT 21291

RESULT 5  
US-10-363-829-150/c  
Sequence 150, Application US/10363829

PUBLICATION No. US2004014231A1

GENERAL INFORMATION:

APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;

APPLICANT: Altus, Christina M.; Dufour, Gerard E.;

APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;

APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;

APPLICANT: Wright, Rachel U.; Gietzen, Darryl;

APPLICANT: Liu, Tommy F.; Yap, Pierre E.;

APPLICANT: Dahl, Christopher R.; Momiya, Monika G.;

APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;

APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;

APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyna H.;

APPLICANT: David, Marie H.; Panzer, Scott R.;

APPLICANT: Flores, Vincent Z.; Daffo, Abel;

APPLICANT: Marwaha, Rakesh; Chen, Alice J.;

APPLICANT: Chang, Simon C.; Au, Alan P.;

APPLICANT: Iman, Rebekah R.

TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT

FILE REFERENCE: PT-1183 USN

CURRENT APPLICATION NUMBER: US/10/363, 829

PRIOR FILING DATE: 2003-03-05

PRIOR APPLICATION NUMBER: PCT/US01/27628

PRIOR FILING DATE: 2001-09-05

PRIOR APPLICATION NUMBER: US 60/229, 751

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 749

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 750

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 747

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 748

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/230, 583

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/230, 517

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: US 60/230, 610

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: US 60/230, 597

PRIOR FILING DATE: 2000-09-06

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 506

SOFTWARE: PERL Program

SEQ ID NO 150

LENGTH: 772

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: LG:944938.1:2000SEP08

US-10-363-829-150

Query Match 22.3%; Score 31; DB 19; Length 772;

Best Local Similarity 57.9%; Pred. No. 7;  
Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 45 TTCTTCGAAAGAGCGGATATTGTCGTGTTATTAAGTATCCCAAGCACC 104  
DB 549 TTCAATTAATAAAGAGGCTCCATTTAGTTATTTATCAGAGTTGTCCAAAAGCGCT 490

QY 105 ATAAATCAAGCTAGCTGTTCTTAATTTGTTAACA 139  
DB 489 TTCTGTATGCTTAGAAATCTCTCACTGCACACA 455

RESULT 6  
US-10-363-829-237/c  
Sequence 237, Application US/10363829

PUBLICATION No. US2004014231A1

GENERAL INFORMATION:

APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;

APPLICANT: Altus, Christina M.; Dufour, Gerard E.;

APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;

APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;

APPLICANT: Wright, Rachel U.; Gietzen, Darryl;

APPLICANT: Liu, Tommy F.; Yap, Pierre E.;

APPLICANT: Dahl, Christopher R.; Momiya, Monika G.;

APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;

APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;

APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyna H.;

APPLICANT: David, Marie H.; Panzer, Scott R.;

APPLICANT: Flores, Vincent Z.; Daffo, Abel;

APPLICANT: Marwaha, Rakesh; Chen, Alice J.;

APPLICANT: Chang, Simon C.; Au, Alan P.;

APPLICANT: Iman, Rebekah R.

TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT

FILE REFERENCE: PT-1183 USN

CURRENT APPLICATION NUMBER: US/10/363, 829

PRIOR FILING DATE: 2003-03-05

PRIOR APPLICATION NUMBER: PCT/US01/27628

PRIOR FILING DATE: 2001-09-05

PRIOR APPLICATION NUMBER: US 60/229, 751

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 749

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 750

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 747

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/229, 748

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/230, 583

PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: US 60/230, 517

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: US 60/230, 610

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: US 60/230, 597

PRIOR FILING DATE: 2000-09-06

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 506

SOFTWARE: PERL Program

SEQ ID NO 237

LENGTH: 810

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: LG:814261.1:2000SEP08

US-10-363-829-237

Query Match 22.3%; Score 31; DB 19; Length 810;  
Best Local Similarity 57.9%; Pred. No. 7.1;  
Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 45 TTCTTCGAAAGAGCGGATATTGTCGTGTTATTAAGTATCCCAAGCACC 104



```

? RESULT 8
? US-09-738-626-2992
? Sequence 2992, Application US/09738626
? Publication No. US20020197605A1
? GENERAL INFORMATION:
? APPLICANT: NAKAGAWA, SATOSHI
? APPLICANT: MIZOGUCHI, HIROSHI
? APPLICANT: ANDO, SEIKO
? APPLICANT: HAYASHI, MIKIRO
? APPLICANT: OCHIAI, KEIRO
? APPLICANT: YOKOI, HARUHIKO
? APPLICANT: TATEISHI, NAOKO
? APPLICANT: SENOH, AKIHIRO
? APPLICANT: IKEDA, MASATO
? APPLICANT: OZAKI, AKIO
? TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
? FILE REFERENCE: 249-125
? CURRENT APPLICATION NUMBER: US/09/738,626
? PRIOR FILING DATE: 2000-12-18
? PRIOR APPLICATION NUMBER: JP 99/377484
? PRIOR FILING DATE: 1999-12-16
? PRIOR APPLICATION NUMBER: JP 00/159162
? PRIOR FILING DATE: 2000-04-07
? PRIOR APPLICATION NUMBER: JP 00/280988
? PRIOR FILING DATE: 2000-08-03
? NUMBER OF SEQ ID NOS: 7059

```

```

Query Match Similarity 22.0%; Score 30.6; DB 9; Length 3309400;
Best Local Similarity 53.8%; Pred. No. 3e+02;
Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 17 AACTAAATTAATGCTGCCTCACCCTCTTTTCTTCAGAAAGAGGATGACTATTGTCTGG 76
Db 2900056 AACGCATTTGATGGACGAGGAGGATGCTTCATGAGAAAGAACTCAACCCATTGGTGGG 2899997

Qy 77 TTTATTAACTGTTTATCCCCAAAGACCATATATCAAGCTGACTGTTCTTATTGTT 133
Db 2899936 ATCATCGACAGTGTCTGTTTCAACCAACCAATTACCGATCCAACGTGATATCGGT 2899940

RESULT 10
US-10-737-082-4/c
; Sequence 4, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris

```



APPLICANT: Burgess, Chris  
APPLICANT: Gannon, Allison  
APPLICANT: Harvey, Jeanne  
APPLICANT: Lechner, John F.  
APPLICANT: Li, Zheng  
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
FILE REFERENCE: 1657/2032  
CURRENT APPLICATION NUMBER: US/10/737,082  
CURRENT FILING DATE: 2003-12-16  
PRIOR APPLICATION NUMBER: US 10/737,082  
NUMBER OF SEQ ID NOS: 300  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 4  
LENGTH: 15649  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-737-082-4

Query Match 21.9%; Score 30.4; DB 22; Length 15649;  
Best Local Similarity 53.3%; Pred. No. 40;  
Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 16 GAACCTTAATAATGCGCTGACCCCTCTTTCTTCAGAAAGGGTGACTATTGCTG 75  
DB 2478 GAACCCAGATTATTCATGGCTAGCTTATGATTAGACTCAGATTCATTCTAG 2419  
QY 76 GTTATTAACTGTTTATCCCAAGCACCATTATCAAGCTAGCTGTTCTTATTGTTAA 135  
DB 2418 GGTCAAGTCTGTTTATTAAGAAATTAAGACAGACCCAGCTGTTCTGATTTCTTAA 2359

RESULT 11  
US-10-765-790-4/c  
Sequence 4, Application US/10765790  
Publication No. US20050130172A1  
GENERAL INFORMATION:  
APPLICANT: Bayer Healthcare LLC  
APPLICANT: Beard, Chris  
APPLICANT: Burgess, Chris  
APPLICANT: Gannon, Allison  
APPLICANT: Harvey, Jeanne  
APPLICANT: Lechner, John F.  
APPLICANT: Li, Zheng  
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
FILE REFERENCE: 1657/2035  
CURRENT APPLICATION NUMBER: US/10/765,790  
CURRENT FILING DATE: 2004-01-27  
PRIOR APPLICATION NUMBER: US 10/737,082  
NUMBER OF SEQ ID NOS: 300  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 4  
LENGTH: 15649  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-765-790-4

Query Match 21.9%; Score 30.4; DB 22; Length 15649;  
Best Local Similarity 53.3%; Pred. No. 40;  
Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 16 GAACCTTAATAATGCGCTGACCCCTCTTTCTTCAGAAAGGGTGACTATTGCTG 75  
DB 2478 GAACCCAGATTATTCATGGCTAGCTTATGATTAGACTCAGATTCATTCTAG 2419  
QY 76 GTTATTAACTGTTTATCCCAAGCACCATTATCAAGCTAGCTGTTCTTATTGTTAA 135  
DB 2418 GGTCAAGTCTGTTTATTAAGAAATTAAGACAGACCCAGCTGTTCTGATTTCTTAA 2359

RESULT 12  
US-09-960-870-13/c

Sequence 13, Application US/09960870  
Publication No. US20030134281A1  
GENERAL INFORMATION:  
APPLICANT: Evans, Glen  
TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF  
FILE REFERENCE: P-EA 4738  
CURRENT APPLICATION NUMBER: US/09/960,870  
CURRENT FILING DATE: 2001-09-20  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 31241  
TYPE: DNA  
ORGANISM: M. genitalium  
US-09-960-870-13

Query Match 21.9%; Score 30.4; DB 10; Length 31241;  
Best Local Similarity 55.8%; Pred. No. 53;  
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 34 CCTGACCCCTCTTTCTTCAGAAAGGGTGACTATTGCTGTTTAACTGTTATC 93  
DB 25866 CCTGACCAATTATTTTAAACCAAGTGGTAGTTTACCATCTGTTGCTCTTCTGTC 25807  
QY 94 CCCAAGCACCATTATCAAGCTAGCTGTTCTTATTGTTTAA 137  
DB 25806 ATTAAAGAACGCTATCAATCTATACCCCTTTTGTGCTAA 25763

RESULT 13  
US-09-960-858-13/c  
Sequence 13, Application US/09960858  
Publication No. US20030138777A1  
GENERAL INFORMATION:  
APPLICANT: Evans, Glen  
TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF  
FILE REFERENCE: P-EA 4974  
CURRENT APPLICATION NUMBER: US/09/960,858  
CURRENT FILING DATE: 2001-09-20  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 31241  
TYPE: DNA  
ORGANISM: M. genitalium  
US-09-960-858-13

Query Match 21.9%; Score 30.4; DB 10; Length 31241;  
Best Local Similarity 55.8%; Pred. No. 53;  
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 34 CCTGACCCCTCTTTCTTCAGAAAGGGTGACTATTGCTGTTTAACTGTTATC 93  
DB 25866 CCTGACCAATTATTTTAAACCAAGTGGTAGTTTACCATCTGTTGCTCTTCTGTC 25807  
QY 94 CCCAAGCACCATTATCAAGCTAGCTGTTCTTATTGTTTAA 137  
DB 25806 ATTAAAGAACGCTATCAATCTATACCCCTTTTGTGCTAA 25763

RESULT 14  
US-10-251-668-13/c  
Sequence 13, Application US/10251668  
Publication No. US20040063097A1  
GENERAL INFORMATION:  
APPLICANT: Evans, Glen  
TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF  
FILE REFERENCE: P-EA 5441  
CURRENT APPLICATION NUMBER: US/10/251,668  
CURRENT FILING DATE: 2002-09-20

```

1 FEATURE:
2 NAME/KEY: misc_feature
3 LOCATION: (362)..(362)
4 OTHER INFORMATION: n is a, c, g, or t
5 FEATURE:
6 NAME/KEY: misc_feature
7 LOCATION: (11659)..(11659)
8 OTHER INFORMATION: n is a, c, g, or t
9 FEATURE:
10 NAME/KEY: misc_feature
11 LOCATION: (11699)..(11699)
12 OTHER INFORMATION: n is a, c, g, or t
13 FEATURE:
14 NAME/KEY: MS4A7_initial_coding_region
15 LOCATION: (11493)..(117539)
16 FEATURE:
17 NAME/KEY: MS4A7_coding_region
18 LOCATION: (19439)..(19573)
19 FEATURE:
20 NAME/KEY: MS4A7_coding_region
21 LOCATION: (21068)..(21124)
22 FEATURE:
23 NAME/KEY: MS4A7_coding_region
24 LOCATION: (23741)..(23947)
25 FEATURE:
26 NAME/KEY: MS4A7_coding_region
27 LOCATION: (27037)..(27138)
28 FEATURE:
29 NAME/KEY: MS4A7_coding_region
30 LOCATION: (28139)..(28210)
31 FEATURE:
32 NAME/KEY: misc_feature
33 LOCATION: (33240)..(32640)
34 OTHER INFORMATION: n is a, c, g, or t
35 FEATURE:
36 NAME/KEY: MS4A5_initial_coding_region
37 LOCATION: (64028)..(64180)
38 FEATURE:
39 NAME/KEY: MS4A5_coding_region
40 LOCATION: (65149)..(65277)
41 FEATURE:
42 NAME/KEY: MS4A5_coding_region
43 LOCATION: (66764)..(66820)
44 FEATURE:
45 NAME/KEY: MS4A5_coding_region
46 LOCATION: (68118)..(68270)
47 FEATURE:
48 NAME/KEY: MS4A5_coding_region
49 LOCATION: (82002)..(82109)
50 FEATURE:
51 NAME/KEY: MS4A12_initial_coding_region
52 LOCATION: (91488)..(91760)
53 FEATURE:
54 NAME/KEY: MS4A12_coding_region
55 LOCATION: (95210)..(95377)
56 FEATURE:
57 NAME/KEY: MS4A12_coding_region
58 LOCATION: (96148)..(96204)
59 FEATURE:
60 NAME/KEY: MS4A12_coding_region
61 LOCATION: (97665)..(97981)
62 FEATURE:
63 NAME/KEY: MS4A12_coding_region
64 LOCATION: (100890)..(101000)
65 FEATURE:
66 NAME/KEY: MS4A12_coding_region
67 LOCATION: (101182)..(101283)
68 FEATURE:
69 NAME/KEY: misc_feature
70 LOCATION: (104561)..(104625)
71 OTHER INFORMATION: n is a, c, g, or t
72 US-10-433-287-79

```

Query Match 21.9%; Score 30.4; DB 19; Length 104644;  
 Best Local Similarity 53.3%; Pred. No. 89;  
 Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
 Qy 16 GAACTTAATAATGCGCTGCTCAACCCTCTTTCTTCAGAAAGGGTGACTATTGTCTG 75  
 Db 88420 GAAACCAGATTATTCATGCTAGTCTATAGTTATAGGACTCAGAAATACATTTTACTG 88361  
 Qy 76 GTTATTAAGTGTATATCCCAAGACATATCAAGCTAGACTGTCTTATTTGTTAA 135  
 Db 88360 GGTCAAGTTCTGTTTATATAAAGAAATTAAGACAGAGCCAGCCAGTTCTGATTCTTTA 88301

Search completed: August 4, 2005, 01:55:06  
 Job time : 363 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:47:30 ; Search time 105 Seconds  
(without alignments)  
2166.119 Million cell updates/sec

Title: US-10-070-882A-2

Perfect score: 139  
Sequence: 1 gtgactctgtgcgcagcaact.....ctgtctctatgttcaacaca 139

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/ECTUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	33	23.7	52494	US-09-949-016-16498	Sequence 16498, A
C 2	32.2	23.2	86273	US-09-949-016-15273	Sequence 15273, A
C 3	31.8	22.9	102304	US-09-949-016-12589	Sequence 12589, A
C 4	30	21.6	601	US-09-949-016-81832	Sequence 81832, A
C 5	30	21.6	601	US-09-949-016-144837	Sequence 144837, A
C 6	30	21.6	14712	US-09-949-016-15840	Sequence 15840, A
C 7	30	21.6	22823	US-09-949-016-14121	Sequence 14121, A
C 8	30	21.6	211049	US-09-949-016-15770	Sequence 15770, A
C 9	29.8	21.4	301	US-09-439-313-293	Sequence 293, App
C 10	29.8	21.4	301	US-09-352-616A-293	Sequence 293, App
C 11	29.8	21.4	301	US-09-232-149A-293	Sequence 293, App
C 12	29.8	21.4	301	US-09-159-812-293	Sequence 293, App
C 13	29.8	21.4	301	US-09-636-215-293	Sequence 293, App
C 14	29.8	21.4	301	US-09-685-166A-293	Sequence 293, App
C 15	29.8	21.4	301	US-09-688-489-293	Sequence 293, App
C 16	29.8	21.4	301	US-09-679-426-293	Sequence 293, App
C 17	29.8	21.4	301	US-09-759-143-293	Sequence 293, App
C 18	29.8	21.4	301	US-09-651-236-293	Sequence 293, App
C 19	29.8	21.4	1140	US-09-107-532A-1865	Sequence 1865, App
C 20	29.8	21.4	2143	US-09-673-395A-82	Sequence 82, App1
C 21	29.8	21.4	3174	US-09-489-847-86	Sequence 86, App1
C 22	29.2	21.0	16593	US-08-961-527-52	Sequence 52, App1
C 23	29	20.9	3010	US-08-714-918-71	Sequence 71, App1
C 24	29	20.9	3010	US-09-265-315-71	Sequence 71, App1
C 25	29	20.9	3010	US-09-265-315-71	Sequence 71, App1
C 26	29	20.9	3010	US-09-266-417-71	Sequence 71, App1
C 27	29	20.9	3010	US-09-528-709-71	Sequence 71, App1

28	29	20.9	3010	4	US-09-527-745-71	Sequence 71, App1
29	29	20.9	194790	4	US-09-949-016-15393	Sequence 15393, A
30	29	20.9	194889	4	US-09-949-016-15654	Sequence 15654, A
C 31	28.8	20.7	601	4	US-09-949-016-47920	Sequence 47920, A
C 32	28.8	20.7	2347	1	US-08-453-695A-113	Sequence 113, App
C 33	28.8	20.7	2347	1	US-08-268-161A-113	Sequence 113, App
C 34	28.8	20.7	2347	2	US-08-453-702A-113	Sequence 113, App
C 35	28.8	20.7	2347	3	US-09-099-639-113	Sequence 113, App
C 36	28.8	20.7	2347	5	PCT-US95-08071-113	Sequence 113, App
C 37	28.6	20.6	601	4	US-09-949-016-40625	Sequence 40625, A
C 38	28.6	20.6	601	4	US-09-949-016-40626	Sequence 40626, A
C 39	28.6	20.6	601	4	US-09-949-016-148424	Sequence 148424, A
C 40	28.6	20.6	601	4	US-09-949-016-148425	Sequence 148425, A
C 41	28.6	20.6	2676	3	US-09-212-971-11	Sequence 11, App1
C 42	28.6	20.6	2676	3	US-08-800-929A-11	Sequence 11, App1
C 43	28.6	20.6	2676	3	US-09-617-053A-11	Sequence 11, App1
C 44	28.6	20.6	16397	4	US-08-956-171E-205	Sequence 205, App
C 45	28.6	20.6	16397	4	US-08-781-986A-205	Sequence 205, App

#### ALIGNMENTS

```

RESULT 1
US-09-949-016-16498/c
; Sequence 16498, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16498
; LENGTH: 52494
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(52494)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16498

Query Match      23.7%; Score 33; DB 4; Length 52494;
Best local similarity 65.8%; Pred. No. 0.66;
Matches 48; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY      18 ACTTAATATATGCTGCTCCTCTTTCTTTCAGAAAGGCTACTATTGTCGT 77
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      11879 ATTGAATATGCTGCTGCTTTCTTCTTTTACCAATATGCTAATGATGCTATT 11820
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      78 TTATTACTGTT 90
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      11819 TTATTAAAGTTT 11807
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-949-016-15273/c
; Sequence 15273, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15273  
LENGTH: 86273  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15273

Query Match 23.2%; Score 32.2; DB 4; Length 86273;  
Best Local Similarity 54.7%; Pred. No. 1.6;  
Matches 64; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 23 AATAAGCCTGCTGACCCCTCTTTCTTTCAGAAAGGGTGACTATTGCTGCTTATT 82  
DB 45531 AATAAGCTGTGACCAAACTATTTCCTGGAAGATTGACAAATTTCTGCTCACT 45472  
QY 83 AACTGTTATCCCAAGCACCATAATCAACCTAGACTGTTCTTATTGTTAACA 139  
DB 45471 TACATTATTTCATTATTAATTTCACAAAGAGAGAGTTGGATTTTAAAAAATA 45415

RESULT 3  
US-09-949-016-12589/c  
Sequence 12589, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12589  
LENGTH: 102304  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(102304)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12589

Query Match 22.9%; Score 31.8; DB 4; Length 102304;  
Best Local Similarity 59.3%; Pred. No. 2.3;  
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 20 TTTAATAATGCTGCTGACCCCTCTTTCTTTCAGAAAGGGTGACTATTGCTGCTTT 79  
DB 6060 TTTAGCACTGCCAACTCATCTCTCTTCCAGAAATTAATAATATTGTCATCTTG 6001  
QY 80 ATTAAGTGTATCCCAAGCACCATAATC 110  
DB 6000 ATTAAGTGTATGACAGAGCAGGAGTATC 5970

RESULT 4

US-09-949-016-81832/c  
Sequence 81832, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 81832  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-81832

Query Match 21.6%; Score 30; DB 4; Length 601;  
Best Local Similarity 54.5%; Pred. No. 1.1;  
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GTGACTGTGTCGACGAATTAATATGCTGCTGACCCCTCTTTCTTTCAGAAAGG 60  
DB 154 GAGACTTTGGGCAATCACTTAATCTGTGCTGCTGCTTCTTATCTATTAATGGG 95  
QY 61 GTGACTATTGTCGTGTTTATTAATGTTATCCCAAGCACCATAATC 110  
DB 94 GAGATTAATGTCGTCACCTTTTACTTTATTAAGATTAAATAGATAGTC 45

RESULT 5  
US-09-949-016-144837/c  
Sequence 144837, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 144837  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-144837

Query Match 21.6%; Score 30; DB 4; Length 601;  
Best Local Similarity 54.5%; Pred. No. 1.1;  
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GTGACTGTGTCGACGAATTAATATGCTGCTGACCCCTCTTTCTTTCAGAAAGG 60  
DB 154 GAGACTTTGGGCAATCACTTAATCTGTGCTGCTGCTTCTTATCTATTAATGGG 95  
QY 61 GTGACTATTGTCGTGTTTATTAATGTTATCCCAAGCACCATAATC 110  
DB 94 GAGATTAATGTCGTCACCTTTTACTTTATTAAGATTAAATAGATAGTC 45

RESULT 6  
US-09-949-016-15840  
; Sequence 15840, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15840  
; LENGTH: 14712  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15840

Query Match 21.6%; Score 30; DB 4; Length 14712;  
Best Local Similarity 54.5%; Pred. No. 4.4;  
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GTGACTTGTGTCGACCAATTAAATATCCCTGCTCACCCTTTCTTCAGAAAGAG 60  
DB 3243 GAGACTTTGGCGAAATCACTTATCTGTCTGCGCTCAGTTTCCTATCTATAAATGG 3302  
QY 61 GTGACTTGTGTCGCTTTTATTAAGTGTATCCCAAGCAACATATC 110  
DB 3303 GAGAAATATGCTGTCACCTTTTACTTTATAAAGATTAAATAGATGTC 3352

RESULT 7

US-09-949-016-14121  
; Sequence 14121, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14121  
; LENGTH: 22823  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14121

Query Match 21.6%; Score 30; DB 4; Length 22823;  
Best Local Similarity 54.5%; Pred. No. 5.3;  
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GTGACTTGTGTCGACCAATTAAATATGCTGCTCACCCTTTCTTCAGAAAGAG 60  
DB 3243 GAGACTTTGGCGAAATCACTTATCTGTCTGCGCTCAGTTTCCTATCTATAAATGG 3302

QY 61 GTGACTTGTGTCGCTTTTATTAAGTGTATCCCAAGCAACATATC 110  
DB 3303 GAGAAATATGCTGTCACCTTTTACTTTATAAAGATTAAATAGATGTC 3352

RESULT 8  
US-09-949-016-15770/c  
; Sequence 15770, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15770  
; LENGTH: 211049  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15770

Query Match 21.6%; Score 30; DB 4; Length 211049;  
Best Local Similarity 61.5%; Pred. No. 14;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 13 GAGCACTTAATTAATATGCTGCTCACCCTTTCTTCAGAAAGAGGCTGACTATTGT 72  
DB 164239 GAGGCCCTTACTTCTTATCTCATCATCCTTTATCTCAATATGAAAGATTCATTTC 164240  
QY 73 CTGCTTATTAAGTGT 90  
DB 164239 ATTGTTGTTAATTATT 164222

RESULT 9

US-09-439-313-293  
; Sequence 293, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqun  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary  
; APPLICANT: Retter, Mark  
; APPLICANT: Solk, John  
; APPLICANT: Day, Craig  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C9  
; CURRENT APPLICATION NUMBER: US/09/439,313  
; CURRENT FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 575  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 293  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-439-313-293

Query Match 21.4%; Score 29.8; DB 3; Length 301;





TYPE: DNA  
ORGANISM: Homo sapien  
US-09-686-215-293

Query Match  
Best Local Similarity 21.4%; Score 29.8; DB 4; Length 301;  
Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 43 CTTTCTTCAGAAAGGGTGAATTTGCTGTTTATTAACGTGTTATCCCAAGCA 102  
DB 30 CTGTTCTCAGTAAAGTCTGGCTAATGCTTGTGTGATCTTGATTCGACATC 89  
QY 103 CCATATCAAGCGCTAGACTGTTCTTATTTGTTAACA 139  
DB 90 AATCAATCAATGGCCTAGAGCACTGACTGTTAACA 126

## RESULT 14

US-09-685-166A-293  
Sequence 293, Application US/09685166A

PATENT NO. 6630305

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, David C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan U.

APPLICANT: Jiang, Yuqun

APPLICANT: Henderson, Robert A.

APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C21

CURRENT APPLICATION NUMBER: US/09/685,166A

CURRENT FILING DATE: 2000-10-10

NUMBER OF SEQ ID NOS: 898

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 293

LENGTH: 301

TYPE: DNA

ORGANISM: Homo sapien

US-09-685-166A-293

Query Match

Best Local Similarity 21.4%; Score 29.8; DB 4; Length 301;

Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 43 CTTTCTTCAGAAAGGGTGAATTTGCTGTTTATTAACGTGTTATCCCAAGCA 102  
DB 30 CTGTTCTCAGTAAAGTCTGGCTAATGCTTGTGTGATCTTGATTCGACATC 89  
QY 103 CCATATCAAGCGCTAGACTGTTCTTATTTGTTAACA 139  
DB 90 AATCAATCAATGGCCTAGAGCACTGACTGTTAACA 126

## RESULT 15

US-09-688-489-293

Sequence 293, Application US/09688489

PATENT NO. 6664377

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, David C.

APPLICANT: Mitcham, Jennifer Lynn

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE

TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.427D2

CURRENT APPLICATION NUMBER: US/09/688,489

CURRENT FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 338

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 293

LENGTH: 301

TYPE: DNA

ORGANISM: Homo sapien

US-09-688-489-293

Query Match

Best Local Similarity 21.4%; Score 29.8; DB 4; Length 301;

Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 43 CTTTCTTCAGAAAGGGTGAATTTGCTGTTTATTAACGTGTTATCCCAAGCA 102  
DB 30 CTGTTCTCAGTAAAGTCTGGCTAATGCTTGTGTGATCTTGATTCGACATC 89  
QY 103 CCATATCAAGCGCTAGACTGTTCTTATTTGTTAACA 139  
DB 90 AATCAATCAATGGCCTAGAGCACTGACTGTTAACA 126

Search completed: August 4, 2005, 01:48:59  
Job time: 107 secs

**This Page Blank (uspto)**



LOCUS STYPHOPA 990 bp DNA linear BCT 26-APR-1993  
 DEFINITION S.typhimurium phop gene encoding Phop virulence protein, complete cds.  
 ACCESSION M25241  
 VERSION M25241.1 GI:154263  
 KEYWORDS phop gene; virulence protein.  
 SOURCE Salmoneilla typhimurium  
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmoneilla.  
 REFERENCE 1 (bases 1 to 990)  
 AUTHORS Groisman,E.A., Chiao,E., Lippe,C.J. and Heffron,F.  
 TITLE Salmoneilla typhimurium phop virulence gene is a transcriptional regulator  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (18), 7077-7081 (1989)  
 MEDLINE 89366683  
 PUBMED 2674945  
 COMMENT Original source text: S.typhimurium (strain 14028a) DNA, clone pEGS381.  
 Draft entry and computer-readable sequence [1] kindly submitted by F.Heffron, 01-JUN-1989.

FEATURES  
 source  
 1..990  
 location/Qualifiers  
 /organism="Salmoneilla typhimurium"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:602"  
 140..142  
 /note="ribosome binding site"  
 151..825  
 /note="Phop protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="AA027187.1"  
 /db\_xref="GI:154264"  
 /translation="MMRVLVVEDNALRRHLKVLQDSGHQVDAEDAREADYLYNEH LPDIAIVDLGDPEDGSLIRMRSSDVS LPVLVLTARGMQDKVEVSSGADYVTK PHIEVYARMQALMRNRGSLASOVINIPPOVDLSRREL SVNEEVILKTAPEYIME TLIRNNGKVVSKDSIMLQVLPDAERESHITDVMGRKKIKIQAGYPHDVITTVAGG YLFELR"

ORIGIN  
 25 minutes on S.typhimurium map.  
 Query Match 100.0%; Score 139; DB 1; Length 990;  
 Best Local Similarity 100.0%; Pred. No. 5,2e-29;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACTCTGGTCGACGAACCTTAATATATGCTGCTCACCCTTTTCTTCAGAAAGAG 60  
 |||||  
 DB 1 GTGACTCTGGTCGACGAACCTTAATATATGCTGCTCACCCTTTTCTTCAGAAAGAG 60  
 |||||  
 QY 61 GTGACTATTTGCTGGTTTATTAAGCTTTATCCCAAGACCATATCAACGCTAGAC 120  
 |||||  
 DB 61 GTGACTATTTGCTGGTTTATTAAGCTTTATCCCAAGACCATATCAACGCTAGAC 120  
 |||||  
 QY 121 TGTTCCTATTGTTAACACA 139  
 |||||  
 DB 121 TGTTCCTATTGTTAACACA 139  
 |||||

RESULT 3  
 STY272210 2459 bp DNA linear BCT 30-NOV-2001  
 LOCUS Salmoneilla typhimurium phop gene and phoQ gene, strain SL1344.  
 DEFINITION AJ272210  
 ACCESSION AJ272210.1 GI:7007368  
 VERSION AJ272210.1 GI:7007368  
 KEYWORDS membrane sensor protein; phop gene; phoQ gene; transcriptional regulator; virulence gene.  
 SOURCE Salmoneilla typhimurium  
 ORGANISM Salmoneilla typhimurium  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmoneilla.  
 REFERENCE 1  
 Cano,D.A., Martinez-Moya,M., Pucciarelli,M.G., Groisman,E.A., Casadesu,J. and Garcia-Del Portillo,F.

TITLE Salmoneilla enterica serovar Typhimurium response involved in attenuation of pathogen intracellular proliferation  
 JOURNAL Infect. Immun. 69 (10), 6463-6474 (2001)  
 MEDLINE 21437654  
 PUBMED 11553591  
 REFERENCE 2 (bases 1 to 2459)  
 AUTHORS Garcia-del Portillo,F.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-FEB-2000) Garcia-del Portillo F., C.S.I.C.-Universidad Autonoma de Madrid, Centro De Biologia Molecular 'Severo Ochoa', Campus de Cantoblanco, 28049 Madrid, SPAIN

FEATURES  
 source  
 1..2459  
 location/Qualifiers  
 /organism="Salmoneilla typhimurium"  
 /mol\_type="genomic DNA"  
 /strain="SL1344"  
 /db\_xref="taxon:602"  
 240..914  
 /gene="phop"  
 240..914  
 /gene="phob"  
 /function="virulence transcriptional regulator"  
 /note="member of a two-component regulatory system"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Phop protein"  
 /protein\_id="CA875591.1"  
 /db\_xref="GI:7007369"  
 /db\_xref="GOA:P14146"  
 /db\_xref="UniProt/Swiss-Prot:P14146"  
 /translation="MMRVLVVEDNALRRHLKVLQDSGHQVDAEDAREADYLYNEH LPDIAIVDLGDPEDGSLIRMRSSDVS LPVLVLTARGMQDKVEVSSGADYVTK PHIEVYARMQALMRNRGSLASOVINIPPOVDLSRREL SVNEEVILKTAPEYIME TLIRNNGKVVSKDSIMLQVLPDAERESHITDVMGRKKIKIQAGYPHDVITTVAGG YLFELR"  
 914..2377  
 /gene="phoQ"  
 914..2377  
 /gene="phoQ"  
 /function="membrane sensor protein"  
 /note="member of a two-component regulatory system"  
 /codon\_start=1  
 /transl\_table=11  
 /product="PhoQ protein"  
 /protein\_id="CA875592.1"  
 /db\_xref="GI:7007370"  
 /db\_xref="GOA:P14147"  
 /db\_xref="UniProt/Swiss-Prot:P14147"  
 /translation="MMRVLVVEDNALRRHLKVLQDSGHQVDAEDAREADYLYNEH LPDIAIVDLGDPEDGSLIRMRSSDVS LPVLVLTARGMQDKVEVSSGADYVTK PHIEVYARMQALMRNRGSLASOVINIPPOVDLSRREL SVNEEVILKTAPEYIME TLIRNNGKVVSKDSIMLQVLPDAERESHITDVMGRKKIKIQAGYPHDVITTVAGG YLFELR"  
 914..2377  
 /gene="phoQ"  
 914..2377  
 /gene="phoQ"  
 /function="membrane sensor protein"  
 /note="member of a two-component regulatory system"  
 /codon\_start=1  
 /transl\_table=11  
 /product="PhoQ protein"  
 /protein\_id="CA875592.1"  
 /db\_xref="GI:7007370"  
 /db\_xref="GOA:P14147"  
 /db\_xref="UniProt/Swiss-Prot:P14147"  
 /translation="MMRVLVVEDNALRRHLKVLQDSGHQVDAEDAREADYLYNEH LPDIAIVDLGDPEDGSLIRMRSSDVS LPVLVLTARGMQDKVEVSSGADYVTK PHIEVYARMQALMRNRGSLASOVINIPPOVDLSRREL SVNEEVILKTAPEYIME TLIRNNGKVVSKDSIMLQVLPDAERESHITDVMGRKKIKIQAGYPHDVITTVAGG YLFELR"  
 914..2377  
 /gene="phoQ"  
 914..2377  
 /gene="phoQ"  
 /function="membrane sensor protein"  
 /note="member of a two-component regulatory system"  
 /codon\_start=1  
 /transl\_table=11  
 /product="PhoQ protein"  
 /protein\_id="CA875592.1"  
 /db\_xref="GI:7007370"  
 /db\_xref="GOA:P14147"  
 /db\_xref="UniProt/Swiss-Prot:P14147"  
 /translation="MMRVLVVEDNALRRHLKVLQDSGHQVDAEDAREADYLYNEH LPDIAIVDLGDPEDGSLIRMRSSDVS LPVLVLTARGMQDKVEVSSGADYVTK PHIEVYARMQALMRNRGSLASOVINIPPOVDLSRREL SVNEEVILKTAPEYIME TLIRNNGKVVSKDSIMLQVLPDAERESHITDVMGRKKIKIQAGYPHDVITTVAGG YLFELR"  
 914..2377  
 /gene="phoQ"  
 914..2377  
 /gene="phoQ"  
 /function="membrane sensor protein"  
 /note="member of a two-component regulatory system"  
 /codon\_start=1  
 /transl\_table=11  
 /product="PhoQ protein"  
 /protein\_id="CA875592.1"  
 /db\_xref="GI:7007370"  
 /db\_xref="GOA:P14147"  
 /db\_xref="UniProt/Swiss-Prot:P14147"  
 /translation="MMRVLVVEDNALRRHLKVLQDSGHQVDAEDAREADYLYNEH LPDIAIVDLGDPEDGSLIRMRSSDVS LPVLVLTARGMQDKVEVSSGADYVTK PHIEVYARMQALMRNRGSLASOVINIPPOVDLSRREL SVNEEVILKTAPEYIME TLIRNNGKVVSKDSIMLQVLPDAERESHITDVMGRKKIKIQAGYPHDVITTVAGG YLFELR"

ORIGIN  
 Query Match 100.0%; Score 139; DB 1; Length 2459;  
 Best Local Similarity 100.0%; Pred. No. 4,5e-29;  
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACTCTGGTCGACGAACCTTAATATATGCTGCTCACCCTTTTCTTCAGAAAGAG 60  
 |||||  
 DB 90 GTGACTCTGGTCGACGAACCTTAATATATGCTGCTCACCCTTTTCTTCAGAAAGAG 149  
 |||||  
 QY 61 GTGACTATTTGCTGGTTTATTAAGCTTTATCCCAAGACCATATCAACGCTAGAC 120  
 |||||  
 DB 150 GTGACTATTTGCTGGTTTATTAAGCTTTATCCCAAGACCATATCAACGCTAGAC 209  
 |||||  
 QY 121 TGTTCCTATTGTTAACACA 139  
 |||||

Db 210 TTTCTTATTGTTAAACACA 228

## RESULT 4

AL627269/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

IAELFEVIDPFAIAQVREALTRTIAELADEFLAIYANMHLDEKRVHGDICRTLRN

ACRPLFEGETELANTIVSXYRDANNTDIALALSAVAALQPCRDITMOYDCKM

ODGLVMDKFTILOSSTPAENYLETYSILKRSFSPNPNIRSLIGAPAGSNPAFH

KODSGYOPFVEMLTDLNSRNPQVARSRIEPLIRKRDYDROKROGRALBELKLEN

LSDGLYKITYKALA

532..1332

/gene="STY1078"

/note="Pfam match to entry PF01433 Peptidase\_M1, Peptidase

family M1, score 245.10, E-value 1.9e-84"

1060..1089

/gene="STY1078"

/note="PS00142 Neutral zinc metallopeptidases,

zinc-binding region signature"

3001..4011

/gene="STY1079"

/note="synonym: PYD"

3001..4011

/gene="STY1079"

/note="Orthologue of E. coli PYD (PYD\_ECOLI); Fasta hit

to PYD\_ECOLI (336 aa), 95% identity in 336 aa overlap"

/codon\_start=1

/transl\_table=11

/product="dihydroorotate dehydrogenase"

/protein\_id="CAD08184.1"

/db\_xref="GI:16502233"

/db\_xref="GOA:082759"

/db\_xref="UniProt/Swiss-Prot:082759"

/translation="MYPFVRKALFQDPERAHERTFQQLRITGTPLEALVROKVP

KPVTCMGLTFKPLGLAAGLDGECITLALGAGSGSLEIGTVTRPQGNKPLRF

LVDAGELINRMGFNNIGVNDLVENYENKAFDIDILGINIGKNDQVPEKNDYL

KVAAYAGYAINISSPNTGRLTQYGLADLDLILAIKKKNDLVQVYHKKYVAVKI

APDICEERLIQVADSLRNHNDIGVATNTTDLRSIVQGNKQCGGSLGRLPLQKST

EIRRLSLEKQLPITIGVGIDSVIARAKIAAGATLVQIISGPIFKPKPLIKELVT

HI"

3136..3990

/gene="STY1079"

/note="Pfam match to entry PF01180 DHODHase,

dihydroorotate dehydrogenase, score 563.50, E-value

1.4e-165"

3238..3297

/gene="STY1079"

/note="PS00912 Dihydroorotate dehydrogenase signature 1"

3874..3936

/gene="STY1079"

/note="PS00912 Dihydroorotate dehydrogenase signature 2"

4174..4719

/gene="STY1080"

4174..4719

/gene="STY1080"

/note="Orthologue of E. coli YCBW\_ECOLI; Fasta hit to

YCBW\_ECOLI (192 aa), 90% identity in 180 aa overlap. Note

discrepancy in position of translational start site."

/codon\_start=1

/transl\_table=11

/product="conserved hypothetical protein"

/protein\_id="CAD08185.1"

/db\_xref="GI:16502234"

/db\_xref="UniProt/TREMBL:082758"

/translation="MMRIKPDNNRWYVDEEDRWMLDLANGMLPFRSFRSKMLTPDA

PCPTGFCVDDALYFSFEKCRDPFELTKQRAELVNLVAIRYLKPOPKSNFPAH

GEMWTPGCGDASVNLSDRAEQVNLVVEFGSNALCLLAQGVVIAAGRTWLGDAIK

IANDRIKPOVHCHSBSLEQAV"

/complement(4716..5825)

/gene="STY1081"

/complement(4716..5825)

/note="Orthologue of E. coli P75863; Fasta hit to P75863

(369 aa), 86% identity in 367 aa overlap"

/codon\_start=1

/transl\_table=11

/product="conserved hypothetical protein"

/protein\_id="CAD08186.1"

/db\_xref="GI:16502235"

```

/db_xref="GOA:Q82757"
/db_xref="UniProt/TREMBL:Q82757"
/translation="MATLSRLFHPHVKSMRGIGLTHALADISGLAPDRIFWTSDDGT
PTLRQPMWVRFTPSPLADGILHTLPAQSSALVAFVFTPTPDADATWGNHFLRAV
PTAIAINMGISCFRSDVQLCWGPOLTRRVRKRNAPVLAGADGYPIALTAEALRDQA
RCAPAGWMEQFRPNLIVGVGAAMEDSNMYRLIGVIFPVVYSGRCIPITYSPKQO
KHPSGEPLATIOAPFTADNGDVDRGQNLIRANSQVIRGVDEVELATLTPAAYGATT
LDPSVTPKHPKPDGSLTIDMGQTPCGNQQVLTLEENQIGRIFPISCRAGICCCRIR
LLEBESVPLKTSAMGGDDTILSCSCVPTALRLN"
complement(4737..4955)
/misc_feature
/note="Pfam match to entry PF00111 fer2, 2Fe-2S
iron-sulfur cluster binding domains, score 55.90, E-value
8.5e-13"
gene
5924..8032
5924..8032
/gene="STY1082"
5924..8032
/gene="STY1082"
/note="Similar to Haemophilus influenzae hypothetical
protein h10116 H10116 SM:YCBY_HAEIN (P43542; P43345) fasta
scores: E() = 0, 60.3% id in 710 aa
Orthologue of E. coli YCBY_ECOLI; Fasta hit to YCBY_ECOLI
(702 aa), 92% identity in 702 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD08187.1"
/db_xref="GI:16502236"
/db_xref="GOA:Q82786"
/db_xref="UniProt/TREMBL:Q82786"
/translation="MNSLPASTARGIEELKTELEKTAGVGGVQGGVHPQDPTSLI
YOSIMSLASRIILPMGCKYSLDILDFGQALINWEIIPKATFAVHPSGLINDTI
RNSQGMKAKYDAITVDAFTRRKYLPRPNVDRESPPDRIYNWLNKRTASILDSDGSH
LRGRDRTGLAPITKETLAALVIMRSGMOGTPLBPMCSGTLILBAAMATDRAGL
HRGHMGFGMAQHDKEAIQEWYAAEQATRRARKLAEYSISFPYSDPAIARIKASRAR
RAGIGELLTFEYKDVDAOLSNPLPKPQYGVILNIPYGERLSDPDALIALHSLGRYMK
NOFGGMNLSLSPASDPLDGLSLOLRADKQFVKNQNPGLDQVQKYNHIAETADSDPATVA
EDVANRLRNKLKLEKMAROEGIECYRLYDADLPENVAVDVYGDMAVIOEAPPTV
DAQKARQRLFDIIAATLTSVYLGIPPKKVLVKTBEQKGRQYQYQMSKSGFLVSEYRNA
RLWMSLTLYLDLTGLFLDHRITARMLEBMSKGDPLNLFYSASVHAELGGARNTT
VDMLNRTYLEMARERLRLNGLSGRAHRLIADCLGMLEAMEBQDLFIPIPTFSNKR
MEESDFVORDHVALMKDLKRLRLKKGTTIWFSSNNKGRFMDLGLAELGLAETITQKT
LSPDFARRQIINCWLIRAA"
/misc_feature
5927..7060
/gene="STY1082"
/note="Pfam match to entry PF01170 UPF0020.
Uncharacterized protein family UPF0020, score 781.30,
E-value 3.7e-231"
5906..6541
/gene="STY1082"
/note="PS01261 Uncharacterized protein family UPF0020
signature"
8045..9952
8045..9952
/gene="STY1083"
8045..9952
/gene="STY1083"
/note="Pasta hit to YBIT_ECOLI (530 aa), 32% identity in
526 aa overlap
Fasta hit to YJUK_ECOLI (554 aa), 34% identity in 524 aa
overlap"
Query Match: 96.5%; Score 134.2; DB 1; Length 254050;
Best Local Similarity 97.8%; Pred. No. 4.7e-28;
Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 GTGACTCTGTCGACGAACCTTAATATATATGCTGACCTCTCTTTCTTCAGAAAGAG 60
|||||
Db 175998 GTGAACTCGTGTGACGAACCTTAATATATATGCTGACCTCTTTCTTCAGAAAGAG 175998
|||||
1 GTGACTAATTTGCTGGTTTATTACTGTTTATCTCCCAAGACCATATATCAACGCTAGAC 120
|||||
Db 175998 GTGACTAATTTGCTGGTTTATTACTGTTTATCTCCCAAGACCATATATCAACGCTAGAC 175998
|||||

```

Cy		121	TGCTTATTTGGTAGACA	139	
D6	175878	IGTCTTATTGTATACACA	175860		
RESULT 5					
AE016839					
LOCUS	AE016839	300029 bp	DNA	linear	BCT 21-MAR-2003
DEFINITION	Salmonella enterica subsp. enterica serovar Typhi Ty2, section 6 of 16 of the complete genome.				
ACCESSION	AE016839	AE014613			
VERSION	AE016839.1	GI:29137536			
KEYWORDS					
SOURCE					
ORGANISM	Salmonella enterica subsp. enterica serovar Typhi Ty2				
REFERENCE	Salmonella enterica subsp. enterica serovar Typhi Ty2 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.				
AUTHORS	1 (bases 1 to 300029)				
	Deng, W., Liou, S.-R., Plunkett III, G., Mayhew, G.F., Rose, D.J., Burland, V., Kodoyanni, V., Schwartz, D.C. and Blattner, F.R.				
TITLE	Comparative Genomics of Salmonella enterica Serovar Typhi Strains Ty2 and CT18				
JOURNAL	J. Bacteriol. 185 (7), 2330-2337 (2003)				
MEDLINE	25313167				
PUBMED	12644504				
REFERENCE	2 (bases 1 to 300029)				
AUTHORS	Deng, W., Liou, S.-R., Plunkett, G. III, Mayhew, G.F., Rose, D.J., Burland, V., Kodoyanni, V., Schwartz, D.C. and Blattner, F.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-SEP-2002) Laboratory of Genetics, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA				
FEATURES	Location/Qualifiers				
SOURCE	1..300029				
	/organism="Salmonella enterica subsp. enterica serovar Typhi Ty2"				
	/mol_type="genomic DNA"				
	/strain="Ty2"				
	/seovar="Typhi"				
	/sub_species="enterica"				
	/db_xref="taxon:209261"				
	245..1336				
gene	/locus_tag="t1459"				
	245..1336				
CDS	/locus_tag="t1459"				
	/note="corresponds to STY1522 from Accession AU513382: Salmonella typhi CT18"				
	/codon_start=1				
	/transl_table=1				
	/product="putative secreted hydrolase"				
	/protein_id="AA069099.1"				
	/db_xref="GI:29137537"				
	/translation="MMKSKAMLFIAIPALPSVMAAQCITLAIODKQDIFHRTLE YMDDLPMLTYTPAGTOFDKKTPGSOGISYOAKYPILATSTTDSDILEGNNS AGLSFERNMIMNADLPPLPASSEYKOAIPLYSGMALAPATVGVEVOAIKEGFNP AD ELHFGRDKSFFHTAFYDKKGSI.VVEVBNGKFHYVNPRFWNTGAPFRHLNLNNK YTGTLTVNRSSGTGGIKVMOPDSGLIADLPSSDTVSRRIRGVYTTPAOKTSIAH DANMTLAHMGRFDRPNRIITVDYMGSGBEGNATRKPVSRYVTWTLSDLTHGEEMVAVG QYNDINTWTMSLSQCFKNADAPVFEKINYG"				
	1382..2173				
gene	/locus_tag="t1460"				
	1382..2173				
CDS	/locus_tag="t1460"				
	/note="corresponds to STY1521 from Accession AU513382: Salmonella typhi CT18"				
	/codon_start=1				
	/transl_table=1				
	/product="putative regulatory protein"				
	/protein_id="AA069100.1"				
	/db_xref="GI:29137538"				
	/translation="METTGCGFSCTVINHYHDRODGMNPKIPMIWGMRVIR EVRERLRKYKTANIILKINDINFPPTFLPPREELAKQGVSAASRALIVIEISG HWIQSGNVGVSDRKHLADSYTIELISTRELVDSCARLAANDVADVQIDEAITY HIRMOAIDNNNVGPYSIDPREFALIASGNRNVLPMRSMLMEORINIPVAQDEQS				

gene  
GDRVILNTLNKOHAIYDAMRQSDASAYEGSLHRLRYRKXIVG"  
complement (2232. .3257)  
/locus\_tag="t1461"  
complement (2232. .3257)  
/locus\_tag="t1461"  
/note="corresponds to STY1520 from Accession AL513382:  
Salmonella typhi CT18"  
/codon\_start=1  
/transl\_table=11  
/product="putative alcohol dehydrogenase"  
/protein\_id="AAO69101.1"  
/db\_xref="GI:29137539"  
/translation="MKKRVAVINKEGVEVTRWVDYHPHKKEVYLIVDAAGICGSDIG  
AFRGTNPATYPRILIGHVAVGIVLQESTGHPDNIKGDRIYDIPYVYCGHCVCSGR  
TNCCENLAVIVAHIDGAEQVETHPAHLIKIKIPNVSENAFLAEPITTLALALHPAN  
VKAGYVAIIIDAGAIIGLMAALSAHAYRATPLIDIVBERLRVQKLGVPVYLAADQ  
VIDAKITGHGTAQVIVTASGANSAIRNTLDIASFGRISFTGMPKROEFTSLPTNLIIT  
YKELIDRSRISAGEPDEALMLSTLEINPDVYSKVNLDEIDPAVKELDRPERYL  
KINAVH"  
complement (3232. .4518)  
/locus\_tag="t1462"  
complement (3232. .4518)  
/locus\_tag="t1462"  
/note="corresponds to STY1519 from Accession AL513382:  
Salmonella typhi CT18"  
/codon\_start=1  
/transl\_table=11  
/product="membrane transport protein"  
/protein\_id="AAO69102.1"  
/db\_xref="GI:29137540"  
/translation="MKRKTGMVLVPLFLFYVTLNVDRSALSTAPRLIEKEGFNAA  
EKGMLFSAPFGIYALFNFDGMSADKQPKVFLIALMSVFGGLGVLGTGLMTLI  
VRLVLEGMEGPEYSAAGNKIINMMISKESASAIAGIFASGPILGANSGVPLGGLALST  
GMRPAGIIFLEGVWVLMVYFIVSDKPTMSKLAPEERIDFENHBDVILISDGRATP  
SIGVYKQKPMWATTLAFESYNIILFPLTFWPSYLNHSLHLDIKESLIVTVPIVYG  
AIGWVLGVCQSVYIRITGMALSRLLIIVGCLAGAAVCVAVSGVSTIGSALTMSV  
SIFLLYLGPIYMAVITQDVHKKDYGVSGAMHGLANIISGILGVLTVGFIVQSSKGD  
YAFYLAGAIIAYISLIVFVFKSGFKANESQSCVH"  
complement (4695. .6251)  
/gene="pqaa"  
/locus\_tag="t1463"  
complement (4695. .6251)  
/gene="pqaa"  
/locus\_tag="t1463"  
/note="corresponds to STY1518 from Accession AL513382:  
Salmonella typhi CT18"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAO69103.1"  
/db\_xref="GI:29137541"  
/translation="MKRKYLMKLCMLAPALVKGESMKIYLVVIVLFFISTKYVTL  
HNINIFCRNSPCDLSHYLPDVRREISGTPAKTYLISAPLAVVAVHVALISQHSMP  
DMVTPAQRNRHNDVYIIPETAKEHVALVANNNGINERGIQIPSKVPDFTQOTASTA  
RQNTNIVSVSDIPQVYLPQDKKPLEDEDSVSVMALFENAEQELMPLNIPWT  
ALSQMARLAKELTQOMINSPFIITGISKRGVTLWSAIDPDEAIVPFIIDLIDIDA  
SLIEHYSQGVFWPIPTFVYQOQIDEKIKSPTFQALQIIDPRVANTVYQPLAP  
KTIYIANGSDPFGVPPNTRPVYSKLPGVSLIIVNMNHSINQFAEGSLVPIFRQPS  
KTLIPOLICLIHMLITVFESEAPKAVYRWANPNARDPRVAGICVQPLTIDIPAN  
NKISTLNSPKGMEATYIETAFENDGVATQSVYITPDEKIPOTAPSPVNAACOTTLFG  
RIGLENDSPD"  
complement (6321. .7562)  
/locus\_tag="t1464"  
complement (6321. .7562)  
/locus\_tag="t1464"  
/note="corresponds to STY1517 from Accession AL513382:  
Salmonella typhi CT18"  
/codon\_start=1  
/transl\_table=11  
/product="putative multidrug efflux protein"  
/protein\_id="AAO69104.1"  
/db\_xref="GI:29137542"  
/translation="NMTNVYENTDSEITTPLNKRRILFVILLVGLVAASTAAVMSVLDP

Query Match	Best Local Similarity	Score	DB 1	Length
Matches 136;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
96.5%;	Score 134.2;	DB 1;	Length 300029;	
1	GTGACTCTGGTTCGACGAACCTTAATTAAGCCCTGCTACCCCTTTCTTCAGAAAGAGG	60		
250187	GTGAATCTGCTGACGAACCTTAATTAAGGCTGCTGACCCCTTTCTTCAGAAAGAGG	250246		
61	GTGACTATTTGTCTGGTTTAACTGTTTATCCCAAGACACATTAATCAACGCTAGAC	120		
250247	GTGACTATTTGTCTGGTTTAACTGTTTATCCCAAGACACATTAATCAACGCTAGAC	250306		

QY 121 TGTCTATTGTTAACA 139  
 |||||  
 DB 250307 TGTCTATTGTTAACA 250325

RESULT 6  
 STYPHOPO  
 LOCUS  
 DEFINITION S.typhimurium phoP protein and membrane protein phoQ genes,  
 complete cds.  
 STYPHOPO 2190 bp DNA linear BCT 26-APR-1993

ACCESSION M24424.1 GI:154265  
 VERSION M24424  
 KEYWORDS membrane protein; phoP gene; phoQ protein.  
 SOURCE Salmonella typhimurium  
 ORGANISM Salmonella typhimurium  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Salmonella.

REFERENCE 1 (bases 1 to 2190)  
 Miller S.I., Kukulski A.M. and Mekalanos J.J.  
 A two-component regulatory system (phoP phoQ) controls Salmonella  
 typhimurium virulence  
 Proc. Natl. Acad. Sci. U.S.A. 86 (13), 5054-5058 (1989)

JOURNAL MEDLINE  
 PUBMED 2544889  
 COMMENT Original source text: S.typhimurium (strain LT2) DNA.  
 Draft entry and printed copy of sequence for [1] kindly provided by  
 S.I. Miller, 01-MAY-1989.

FEATURES  
 source  
 1..2190  
 Location/Qualifiers

/organism="Salmonella typhimurium"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:602"  
 141..815  
 /note="phoP protein"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="AAA27188.1"  
 /db\_xref="GI:154266"

/translation="MRRVYVVDNALLRHLKYQUDSGHVDAAEDAREADYINEN  
 LPDIAIVDGLPDEGLSLIRMRSDVSLPVLVLRGWDKVEVLSSGADYVTK  
 PHIEVVARQALMRNSGLASQVINIPFQVDLSRELISNEVILKTAETVIME  
 TLRNNGKVVSKDSIMLQLYPAELRESHTIDVLMGRLEKKIQADYPHDVITTVAGCG  
 YVFEELR"  
 815..2170  
 /note="membrane protein phoQ"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="AAA27189.1"  
 /db\_xref="GI:154267"

/translation="MNRKFAHPLPLSLRVRPFLATAGVYVLAAGVATVAVGVSF  
 DKTRFLRGLSMLFTYLLAKMNNKISVLEPENLDNRITPILISIOPEMKITNGFH  
 ELITVNDATSLISEDSHQAQEKIKEYREDDDAEMTHSAVNIYPATAMPELTVV  
 DTPLEKRSYVWVWFVYVLANLVLPLMIAMWLRPIELAREVRELEBHR  
 BMNPETREILSLVNLNLQKSERERVTKRTYLLDTLSLKTPLAVLOSTLESLR  
 NEMSVSKAPVLEQISRIISOIGYILHRASMRSGVLSLELHPVAPLNLISAL  
 NRYORKGVNISMDISPEISPGEQNDPVENMGVLDNAKCYCLREVERISATODHL  
 HIFVEDGPGIIPHSKRSIVFDKGRADITLRPQIILSDSLGGRARHEVVFGRHPYQKE  
 E"

ORIGIN

Query Match 92.8%; Score 129; DB 1; Length 2190;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-26;  
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TCGACCACTTAATATATGCTGCTCACCCTCTTTCTTCAGAAAGAGGCTGACTATT 70  
 |||||  
 DB 1 TCGACCACTTAATATATGCTGCTCACCCTCTTTCTTCAGAAAGAGGCTGACTATT 60  
 |||||  
 QY 71 GTCGTATTATTACGTTTATCCCAAGACCATTAATACGCTAGACTGTTCTTATT 130  
 |||||  
 DB 61 GTCGTATTATTACGTTTATCCCAAGACCATTAATACGCTAGACTGTTCTTATT 120  
 |||||

QY 131 GTTAACACA 139  
 |||||  
 DB 121 GTTAACACA 129

RESULT 7  
 AE008754/c  
 LOCUS  
 DEFINITION Salmomella typhimurium LT2, section 58 of 220 of the complete  
 genome.  
 Salmomella typhimurium LT2, section 58 of 220 of the complete  
 genome.  
 AE008754 AE006468  
 AE008754.1 GI:16419750

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Salmonella typhimurium LT2  
 Salmonella typhimurium LT2  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Salmonella.  
 1 (bases 1 to 24578)  
 McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W.,  
 Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F.,  
 Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A.,  
 Grewal, N., Mulvaney, E., Ryan, R., Sun, H., Flores, L., Miller, W.,  
 Stoneking, T., Nhan, M., Waterston, R., and Wilson, R.K.  
 Complete genome sequence of Salmonella enterica serovar Typhimurium  
 LT2

JOURNAL Nature 413 (6858), 852-856 (2001)  
 MEDLINE 21534948  
 PUBMED 11677609  
 2 (bases 1 to 24578)

REFERENCE  
 AUTHORS  
 TITLES  
 JOURNAL

COMMENT The Salmonella typhimurium Genome Sequencing Project  
 Direct Submission  
 Submitted (29-MAR-2001) Genome Sequencing Center, Department of  
 Genetics, Washington University School of Medicine, 4444 Forest  
 Park Boulevard, St. Louis, MO 63108, USA  
 COMMENT Supported by NIH grant 5U 01 AI43283

COMMENT Coding sequences below are predicted from manually evaluated  
 computer analysis, using similarity information and the programs;  
 GLIMMER: <http://www.tigr.org/softlab/glimmer/glimmer.html> and  
 GeneMark; <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto  
 Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>,  
 and Pedro Romero and Peter Karp at EcoCyc;  
<http://ecocyc.org/PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites  
 were kindly provided by Heladia Salgado, Julio Collado-Vides and  
 ReguonB;  
[http://kinch.cifn.unam.mx:8850/db/reguondb\\_intro.framesec](http://kinch.cifn.unam.mx:8850/db/reguondb_intro.framesec)

This sequence was finished as follows unless otherwise noted: all  
 regions were double stranded, sequenced with an alternate  
 chemistries or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one ml3 subclone.

FEATURES  
 source

Location/Qualifiers  
 1..24578  
 /organism="Salmonella typhimurium LT2"  
 /mol\_type="genomic DNA"  
 /strain="LT2; SGC 1412; ATCC 700720"  
 /db\_xref="ATCC:700720"  
 /db\_xref="taxon:99287"  
 /note="LT2"  
 complement(87..1472)  
 /gene="purB"  
 /note="synonym: STM1232"  
 complement(87..1457)  
 /gene="purB"  
 /EC number="4.3.2.2"  
 /note="similar to E. coli adenylsuccinate lyase

CDS



(AAC74215.1); Blastp hit to AAC74215.1 (456 aa), 94%  
 identity in aa 1 - 456"  
 /codon\_start=1  
 /transl\_table=11  
 /product="adenylosuccinate lyase"  
 /protein\_id="AAU20161.1"  
 /db\_xref="GI:16419751"  
 /translation="MELSLTAVSPVDRYGDKVALRGISEVGLTKRVOVEVRML  
 OKLAHAHAIKEVPAPADANGTLDTLVANFNEEDARIKTERTNHDKAVEFLKE  
 KVAALPALHDVSEFTHFACTSEDINNLSHALMLKTARDEVLIPTRQVINAVIDLATO  
 VRDIPLSRTHGQPATPSTLGEMANVAVRMEROPROANOVELIGKINGAVGVNAHI  
 AAVEPWKHOFSESEVTSIGIOMVYTTQIEPHDYIAELFDICARFNTLLIDPFDVGL  
 GYILNHFROKTIAGEIGSGSTWPKHVNPIDEPNSGDLGSAVLHNLANKLPRDWQ  
 RDLDTSTYLRNIGVIGIALLIAYOSTIKGVSLTEVNRDHLDELHNNEVLAEPICVY  
 MRRGIEKPEYELKELTRGKRVDAEGMKQFIDSLALPEAEKRLKAMPANTIGRAVY  
 LVDELK"  
 /complement(1461..2160)  
 /gene="ycfC"  
 /note="synonym: STM1233"  
 /complement(1461..2108)  
 /gene="ycfC"  
 /note="similar to E. coli orf, hypothetical protein  
 (AAC74216.1); Blastp hit to AAC74216.1 (213 aa), 94%  
 identity in aa 1 - 213"  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown membrane associated protein"  
 /protein\_id="AAU20162.1"  
 /db\_xref="GI:16419752"  
 /translation="MIVAKNYDITLALSGICOSARLVQQLAHQHCADADALHSINS  
 VIDMNPSTLGVFGSGSEANLRIGLETLGAVINASSRQGINAELTRYTSLIMLEKLS  
 SKKGLNLTIGDRINGLQROLDHFDLOSDTLKNAMGIYDVISPGRPRQVTSRAVL  
 OSPVOAKVRASLILAGIRAAVLMHVGGRQLMFSRRLTTQAKQIILHLTPEL"  
 /complement(1467..1472)  
 /gene="ycfC"  
 /note="putative RBS for purB; RegulonDB: STMS1H001488"  
 /complement(2112..2117)  
 /gene="ycfC"  
 /note="putative RBS for ycfC; RegulonDB: STMS1H001489"  
 /complement(2134..2142)  
 /gene="ycfC"  
 /note="putative -10 signal for ycfC;  
 RegulonDB: STMLTH004479"  
 /complement(2152..2160)  
 /gene="ycfC"  
 /note="putative -35 signal for ycfC;  
 RegulonDB: STMLTH004479"  
 /complement(2189..3349)  
 /gene="trmU"  
 /note="synonym: STM1234"  
 /complement(2189..3340)  
 /gene="trmU"  
 /BC\_number="2.1.1.61"  
 /note="similar to E. coli orf, hypothetical protein  
 (AAC74217.1); Blastp hit to AAC74217.1 (383 aa), 94%  
 identity in aa 1 - 383"  
 /codon\_start=1  
 /transl\_table=11  
 /product="trmU  
 (5-methylaminomethyl-2-thiouridylylate)-methyltransferase"  
 /protein\_id="AAU20163.1"  
 /db\_xref="GI:16419753"  
 /translation="MIEYALAKFNVASIPWSESPKRVIVGMSGVDSSVASMLLQOOG  
 YVYEGFLPMKNWEEDGEBYCTAAADLAQAVCDKLGIELHTVNPFAEYWDVNFELFL  
 EBYKARTRPMDPLCKEIKERKAFLEPAEDIGADVIATGVRRAVDNKGSRLLFGL  
 DGNKDSYFLTYTGHECIIASLFPVGELEPQVRYKLABDLGLVYATKKGSTGCTFGE  
 RKPRDLGRLPAQPKKITVDDDELGEHQGLMHTLGGKGLGIGTGXDSGSDPWTY  
 VDDQVNNVLIYAQGHHPRLMSVGLIAQGLHWVDRPPTGTLRCTVKKRYRQDTIPC  
 TIALNADNRLEIVFDEPVAAVTPGQSAVFYGSVCIGGGIIEGRPLTV"  
 /complement(3344..3349)  
 /gene="trmU"  
 /note="putative RBS for trmU; RegulonDB: STMS1H001490"  
 /complement(3349..3820)

/gene="ymfB"  
 /note="synonym: STM1235"  
 /complement(3349..3810)  
 /gene="ymfB"  
 /note="similar to E. coli putative phosphohydrolase  
 (AAC74218.1); Blastp hit to AAC74218.1 (153 aa), 84%  
 identity in aa 1 - 152"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative Mult-Like protein"  
 /protein\_id="AAU20164.1"  
 /db\_xref="GI:16419754"  
 /translation="MFKEHVTVAIVHAEDKFLVVEETINGSLSMNQAPAGLEADETL  
 AQAARELMESTGTAQPOVFIKHMOWIAPDKTPELRFLPAIKLDHICATDPHNDID  
 CCHVASEDEINAINLRSPLVAESVRSVQTQORVLSIVGERNPFTEGVK"  
 /complement(3815..3820)  
 /gene="ymfB"  
 /note="putative RBS for ymfB; RegulonDB: STMS1H001491"  
 /complement(3822..4159)  
 /gene="STM1236"  
 /complement(3822..4151)  
 /gene="STM1236"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative periplasmic protein"  
 /protein\_id="AAU20165.1"  
 /db\_xref="GI:16419755"  
 /translation="WKRLILIPALLTPLSVNARSFDCQKAQADEKATCAHLTLNDKD  
 VEMTKYQFLKGLFAMGSRGALQDAQGSWLKORQCKADVTCLTRAVERLKLQLDVY  
 DHIDKPL"  
 /complement(4148..4813)  
 /gene="ymfC"  
 /note="synonym: STM1237"  
 /complement(4146..4813)  
 /gene="ymfC"  
 /note="similar to E. coli orf, hypothetical protein  
 (AAC74219.1); Blastp hit to AAC74219.1 (207 aa), 83%  
 identity in aa 1 - 205"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative ribosomal large subunit pseudouridine  
 synthase"  
 /protein\_id="AAU20166.1"  
 /db\_xref="GI:16419756"  
 /translation="WRQLISSENTMOKTSFRNHYVYKRESSROASRKENQKRVYLF  
 KNPYVVLDPQTEAGRTLKDFIPVQGVYAAERLDSDSGLLVINDGALQARLTPG  
 KTKGRIYQVEGIDPNAQLALRGVTLNDGPPLPAGIEIYAEPMWMPRTPIER  
 KNIPSWLKVTLYEGRRNQVRMTAVGHPTLRLIRYSMDGYTLNGLDNGQREIAQ  
 KDR"  
 /complement(4154..4159)  
 /gene="STM1236"  
 /note="putative RBS for STM1236; RegulonDB: STMS1H001492"  
 /complement(4740..4754)  
 /note="putative binding site for FnrR; RegulonDB:  
 STMS1H00164"  
 /bound\_moiety="FnrR"  
 /gene="icdA"  
 /note="synonym: STM1238"  
 /complement(4829..4837)  
 /gene="icdA"

Query Match 81.3%; Score 113; DB 1; Length 24578;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-22;  
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACTCTGTGACGCAACTTAATAATGCGCCGCTCACCCCTTTCTTCAGAAAGAG 60  
 |||||  
 Db 113 GTGACTCTGTGACGCAACTTAATAATGCGCCGCTCACCCCTTTCTTCAGAAAGAG 54  
 |||||  
 Qy 61 GTGACTATTGTGCTGTTATTAACTGTTATCCCAAGACCATATCAAC 113  
 |||||  
 Db 53 GTGACTATTGTGCTGTTATTAACTGTTATCCCAAGACCATATCAAC 1

RESULT 8  
AE008753/c

LOCUS 22182 bp DNA linear BCT 23-APR-2003

DEFINITION *Salmonella typhimurium* LT2, section 57 of 220 of the complete genome.

ACCESSION AE008753 AE006468

VERSION AE008753.1 GI:16419731

KEYWORDS

SOURCE *Salmonella typhimurium* LT2

ORGANISM *Salmonella typhimurium* LT2

REFERENCE 1 (bases 1 to 22182)  
McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Portolillo, S., All, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stokeling, T., Nhan, M., Waterston, R. and Wilson, R.K.  
Complete genome sequence of *Salmonella enterica* serovar *Typhimurium* LT2

TITLE

JOURNAL Nature 413 (6858), 852-856 (2001)

MEDLINE 21534948

PUBMED 11677609

REFERENCE 2 (bases 1 to 22182)

AUTHORS

CONSRM

TITLE The *Salmonella typhimurium* Genome Sequencing Project

JOURNAL Direct Submission

COMMENT Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA

COMMENT Supported by NIH grant 5U 01 AI43283

COMMENT

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc; <http://ecocyc.pangeasytems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and RegionDB; [http://kinich.cifn.unam.mx:8850/db/regulondb\\_intro.frameset](http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset)

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m33 subclone.

FEATURES

source

1..22182 location/Qualifiers

/organism="Salmonella typhimurium LT2"

/mol\_type="genomic DNA"

/strain="LT2; SGSC 1412; ATCC 700720"

/db\_xref="ATCC:700720"

/db\_xref="taxon:99287"

/note="LT2"

138..407

/gene="ycfR"

/note="synonym: STM1214"

138..143

/gene="ycfR"

/note="putative RBS for ycfR, RegulonDB:STMS1H001474"

150..407

/gene="ycfR"

/note="similar to E. coli orf, hypothetical protein (AAC74196.1); Blastp hit to AAC74196.1 (85 aa), 91%

CDS

identity in aa 1 - 85"

/codon\_start=1

/transl\_table=1

/product="putative outer membrane protein"

/protein\_id="AAI20143.1"

/db\_xref="GI:16419732"

/translation="MKVKVTLIAAVSLSPASFAVEVQATPEGQKFGTISANGG  
TNLSLEPOLAKQAKQAKSGFRITSTVGPVTLHGTAIVYK"

complement(504..1483)

/gene="ycfS"

/note="synonym: STM1215"

complement(504..1469)

/gene="ycfS"

/note="similar to E. coli orf, hypothetical protein (AAC74197.1); Blastp hit to AAC74197.1 (320 aa), 84%

identity in aa 1 - 320"

/codon\_start=1

/transl\_table=1

/product="putative periplasmic protein"

/protein\_id="AAI20144.1"

/db\_xref="GI:16419733"

/translation="WFNNTRLSRWALATLTATMLAPQAQNTPLPPGSRIVGEN  
KFHVEDDGSLEAIKKNVGFLLDQANFGIDPYVRASVLTPIQLTLPDPRS  
GIVNLAELRLVYYPGKNTVTVPIGQLGDTLPETWTTLSDKLANPTPTAN  
IRARYKAQIDLPVAVPAGPDNPMGHAIIRLAAYGVYLLHGYNADFGIKRVSQCI  
RLRODITFLPROVPGTKVNIINTPIKASVEGVRLVEVHQPSSKNIGDDPQVLP  
VLNPMQAFKXAPQTDAAVMEHVEVRSQVVDVTRQAKRQSL"

complement(1478..1483)

/gene="ycfS"

/note="putative RBS for ycfS; RegulonDB:STMS1H001475"

complement(1617..5081)

/gene="mfd"

/note="synonym: STM1216"

complement(1617..5063)

/gene="mfd"

/note="similar to E. coli transcription-repair coupling factor; mutation frequency decline (AAC74198.1); Blastp hit to AAC74198.1 (1148 aa), 95% identity in aa 1 - 1148"

/codon\_start=1

/transl\_table=1

/product="transcription-repair coupling factor"

/protein\_id="AAI20145.1"

/db\_xref="GI:16419734"

/translation="MPKORYTLPTKADQROQLGTLGACATLVAAIEBRAGPVV  
LAPMOMALRLHDEIRORTDOMNGLAMWELTPDSFPHOETISRLSTLYQLEMG  
RGVLIVNTIMQKVCPSYHGHALVWKQKQRLSRDLRDLRQAGRYHDQWSE  
EYATRGALDLDPKMSBOPIRLDFDDSIDSLRLPADTQRTLEVEANLILPAHEP  
TDKAIIEIFRSQWRTFEVKDAEHIYQVQKGTLPAGIEWQPLFESBPLPYSYF  
PANTLVNTGSLSTSAERFOADTLARFENRGVDPMRPLPPEALMLRDELFSELKRW  
PRLQIKTDLPEKANTNLGFOKLPLDIAIQOQKAPLALKRFLFSFGAPVFSVES  
GRRELGLARIXIAPKRIIRLDBAQAQGVYIMGAHEGIDTQRNLATICESDL  
GERVARRILDSRTINPTLIRNLAELHVGQPVYHLENGVRKAMTILERAGITGEY  
LMLTVANAKLYVPSSSLHLSRAGGAEBAPLHLGLGDAMSRARQAEEVRVAA  
ELIIDIYQARAKEGFAFGHDEBOYLFCDSPPFETPPQQAANVNLVDFMCCPLAMR  
LVSCGVGKTEVMAFAFLAVENHGVAVLPTTLAQOAYDNPRDFAMPVATIM  
LSRPSAKGEOQLAABAEGKIDILIGHKLQSVKRDGLLVDESHRGVAKES  
RIKMRADVDTLTATPIRPTIMNAGSGMDLTIAIPAPRAIRAGHQAERFVRYSDLV  
REAIRREILRGQVYLYLVNVDENIQKAERLAELVPEAKIRAIHQKQERREIRAMND  
FHHORENVLCCTTIIETGIDIPYANTIIIEBADHGLQLHQLRQVRSSHQAAML  
LTPHKKMTTDAQKRLAIVASLEDIGAPALATHDLEIRGAGELIGESQSGMETYIGF  
SLYVELLENVADALAKAGREPSLEDLTSGQTEVEMLPDLDDFIPDVNTKFSFKRI  
ASAKNENLEIEIKVELIDRFGLPDPANLIDRLRQAOQGLGIRKLENGKGTIS  
FAEKHNDPAVLIGLQKQPHFRDLDSRTLRKFTIDLSERTRIDWVQFQGLENA  
IA"

complement(5076..5081)

/gene="mfd"

/note="putative RBS for mfd; RegulonDB:STMS1H001476"

5290..6600

/gene="ycfU"

/note="synonym: STM1217"

5290..6600

/gene="ycfU"

/note="similar to E. coli orf, hypothetical protein

CDS

```
(AAC74200.1); Blastp hit to AAC74200.1 (399 aa), 93% identity in aa 1 - 399"
```

```
/codon_start=1  
/transl_table=11  
/product="Integral membrane protein ABC transporter"  
/protein_id="AAU20146.1"  
/db_xref="GI:16419735"  
/translation="MSAFPRITLTNSYSGDIYAFRRLTRPFANSNQTDNYQPVALFIRLAKMGKRAADRPKPSPVSLTIGTLGVALLTVLSVMGFEELDNINILGWFOALTALEHSDNENPNOMPEKAIVNGVRNLAPLTGDVVQLSASVAAGWMLGIDPAKDPLFTYLAVVKOSELDPGRKKVLISGLAQGLVNNGDQRILRWVPASQETPMRLLSQRLTYLGTFEFAENSVDGEEMLVNIODASLMRYPGAINIGRMIMDEPDLVTISOOTLPQGMDWRKERKEIQAAMKENMMGGILLSTLYAIPAFLITSILGMWEKOEVAILIQGTGITPIRMIVPVQQGSACIIIGALLGAALAQSUNMIIPITIGAFLDGAA LVAIAEPQVIATVAAAIALNSTUYTPSRAPAATOAEALKRIE"  
/gene="ycfV"  
/note="synonym: STM1218"  
6580..7294  
5880..6585  
/gene="ycfV"  
/note="putative RBS for ycfV; RegulonDB:STM5IH001477"  
6593..7294  
/gene="ycfV"  
/note="similar to E. coli putative ATP-binding component of a transport system (AAC74201.1) ; Blastp hit to AAC74201.1 (228 aa), 94% identity in aa 1 - 228"  
/codon_start=1  
/transl_table=11  
/product="ATP-binding protein ABC transporter"  
/protein_id="AAU20147.1"  
/db_xref="GI:16419736"  
/translation="NMWKILTCDNDCKRYOEONPDVLDHVSFSIGEGBMAYVGSS GSKGSTLTLLGSLDTPTSGDYFSCQPMKSASSAKAKELRNKGLFIYYPHLLPDF TALLEVAMPULLGDKKKPAIIDARERMLHAVGLEHRATRPSLSGGERVARLARL VNMPHLVLADEPTGNINDANNADSIFELLGELRNLOQTAFVLVTHIDLARMRSOLEM RDGRRTAESIKTMCAE"  
7280..8538  
7280..7285  
/gene="ycfW"  
/note="synonym: STM1219"  
7280..7285  
/gene="ycfV"  
/note="putative RBS for ycfW; RegulonDB:STM5IH001478"  
7294..8538  
/gene="ycfW"  
/note="similar to E. coli putative kinase (AAC74202.1) ; Blastp hit to AAC74202.1 (414 aa), 92% identity in aa 1 - 414"  
/codon_start=1  
/transl_table=11  
/product="integral membrane protein ABC transporter"  
/protein_id="AAU20148.1"  
/db_xref="GI:16419737"  
/translation="MASPPLSLGLGRFSRRRGWGVSLSISTIGIALGVAVLIYG ILSANGFESELENLNRIAIVAHBIEAVNOPWTMRREALFAKVGPGLAAPARYNFNPGL IVESGANULAIQVGDVPOEOOLSAISPVOHAMDHFKAGEEQILIIGGVADVDAFV POGKWYSWIIPANAHDHKILOPERVHTYIGILOGLDRHSFMILEDAOYTLDMSG SSVCIGLATKHVDYPANKKYRDNRGETNSVYTKSYTGTCGYAWRIDQMILAITYL LVIVGACNVNISVILMAVXDSGDIADVLTTEAKOSGLPAIFWVGLAGLOSILIG
```

```

Query Match      61.9% ; Score 86 ; DB 1 ; Length 22182 ;
Best Local Similarity 100.0% ; Pred. No. 3,5e-14 ;
Matches 86 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

QY      54  AAAGAGGAGTGCATATTTGTCGTGTTATTAACGTGTTATCCCAAGACACATATCAAC 113
DB      22182  AAAAGAGGAGTGCATATTTGTCGTGTTATTAACGTGTTATCCCAAGACACATATCAAC 22122

QY      114  GCTAGACTGTCTTATTTGTTAAGACA 139
DB      22122  GCTAGACTGTCTTATTTGTTAAGACA 22097

RESULT 9

```

	LOCUS	AX417848	41 bp	DNA	linear	PAT 18-JUN-2002
	DEFINITION	Sequence 17 from Patent WO0230457.				
	ACCESSION	AX417848				
	VERSION	AX417848.1	GI:21522963			
	KEYWORDS					
	SOURCE					
	ORGANISM	Escherichia coli Escherichia coli Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.				
	REFERENCE	1				
	AUTHORS	Cutlass III, R. and Tinge, S.A.				
	TITLE	Microbes having an attenuating mutation comprising a transcription terminator				
JOURNAL	Patent: WO 0230457-A 17 18-APR-2002;					
FEATURES	MEGAN HEALTH INC (US); WASHINGTON UNIVERSITY IN ST LOUIS (US)					
source	location/Qualifiers					
	1..41					
	/organism="Escherichia coli"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:562"					
ORIGIN						
Query Match	28.8%; Score 40; DB 6; Length 41;					
Best Local Similarity	100.0%; Pred. No. 1.2;					
Matches	40; Conservative 0; Mismatches 0; Indels 0; Gaps 0.					
QY	11	TCGACGACTTAATAATGCGCTGCCTCACCCCTTTTCTT	50			
DB	1	TCGACGACTTAATAATGCGCTGCCTCACCCCTTTTCTT	40			

RESULT 10	
AC110381/c	
LOCUS	166554 bp DNA linear ROD 10-JUL-2004
DEFINITION	Mus musculus BAC clone RP24-19B116 from chromosome 10, complete sequence.
ACCESSION	AC110381
VERSION	AC110381.5
KEYWORDS	GI:49458021
SOURCE	HTS.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 166554) Nguyen,C., Meyer,R., Dignan,G., Creason,K., Cotton,M. and Haglund,K. The sequence of Mus musculus BAC clone RP24-19B116 Unpublished (2001) 2 (bases 1 to 166554) McPherson,J.D. and Waterston,R.H. Direct Submission Submitted (11-FEB-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 3 (bases 1 to 166554) Wilson,R.K. Direct Submission Submitted (08-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 4 (bases 1 to 166554) Wilson,R.K. Direct Submission Submitted (30-JUN-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 5 (bases 1 to 166554) Wilson,R.K. Direct Submission Submitted (10-JUL-2004) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA On Jun 30, 2004 this sequence version replaced gi:47084610.
COMMENT	----- Genome Center Center: Washington University Genome Sequencing Center

Center code: MUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
----- Summary Statistics -----  
Center project name: M\_BB0198116  
-----

## NOTICE:

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

## SOURCE INFORMATION:

The RPCT-24 BAC Library has been constructed by Pieter de Jong and  
coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen  
and/or brain genomic DNA. The clone and detailed information can be  
obtained from Pieter de Jong and coworkers at <http://www.chori.org>

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Location/Qualifiers

1.166554

/organism="Mus musculus"

/mol\_type="Genomic DNA"

/db\_xref="taxon:10090"

/chromosome="10"

/map="10"

/clone="RP24-198116"

/clone\_1lb="RPCT-24"

888.940

/rpt\_family="ID"

1061.1435

/note="Unresolved simple sequence repeat."

1453.1595

/rpt\_family="Alu"

1713.1904

/rpt\_family="B2"

2261.2337

/rpt\_family="ID"

2513.2639

/rpt\_family="Alu"

3903.3993

/rpt\_family="Alu"

5002.5190

/rpt\_family="B2"

5796.5852

/rpt\_family="B2"

5853.6041

/note="Unresolved simple sequence repeat."

6043.6115

/rpt\_family="Alu"

6772.6963

/rpt\_family="B2"

6983.7246

/rpt\_family="MER1\_type"

7369.7521

/rpt\_family="MaLR"

8649.8816

/rpt\_family="MER1\_type"

9075.9188

/rpt\_family="Alu"

9107.9257

/rpt\_family="MaLR"

9283.9468

/rpt\_family="ERVK"

9717.9785

/rpt\_family="ERVK"

9848.9876

/rpt\_family="L2"

10164.10230

/rpt\_family="L1"

10513.10595

/note="Sequence derived from one plasmid subclone."

10547.10685

/rpt\_family="L2"

10691.10832

/rpt\_family="Alu"

10935.11093

/rpt\_family="L1"

11222.11456

/rpt\_family="MER1\_type"

11468.11602

/rpt\_family="B2"

11549.11813

/note="Unresolved tandem repeat."

11604.11852

/rpt\_family="ERV1"

11648.11710

/rpt\_family="ERV1"

11686.11760

/rpt\_family="ERV1"

11754.11810

/rpt\_family="ERV1"

11884.12147

/rpt\_family="L1"

12172.12378

/rpt\_family="L1"

12427.12542

/rpt\_family="Alu"

12559.12690

/rpt\_family="L1"

13153.13354

/rpt\_family="B2"

13639.13968

/rpt\_family="L1"

13969.14311

/rpt\_family="L1"

15247.15367

/rpt\_family="Alu"

16734.16796

/rpt\_family="B4"

18600.18806

/rpt\_family="B2"

19076.19282

/rpt\_family="MER1\_type"

19577.19916

/rpt\_family="MaLR"

21650.21932

/rpt\_family="B4"

21936.22042

/rpt\_family="Alu"

22106.22389

/rpt\_family="MaLR"

23099.23401

/rpt\_family="MaLR"

23635.23898

/rpt\_family="MER1\_type"

24718.25099

/rpt\_family="MaLR"

25134.25197

/rpt\_family="ID"

25386.25462

/rpt\_family="ID"

25565.25678

/rpt\_family="MaLR"

```

repeat_region      25804..25937
                    /rpt_family="L1"
repeat_region      25953..26185
                    /rpt_family="L1"
repeat_region      26439..26532
                    /rpt_family="S8"
repeat_region      26577..26825
                    /rpt_family="B4"
repeat_region      26956..27111
                    /rpt_family="MIR"
repeat_region      27942..28137
                    /rpt_family="B2"
repeat_region      28500..28615
                    /rpt_family="ERV1"
repeat_region      30190..30292

Query Match
Beet Local Similarity 57.4%; Score 36.6; DB 10; Length 16554;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy      2  TGACTCTGCTCGACGAACTTAATAATAGCTGCTCCTCCTCTTTCTTGAGAAAGAGG 61
Db      100283 TGTACTGCTTACAGGCTGTGATGATGCTGCATCATCCTCTCTTCACTAGAGCTG 100224

Qy      62  TGACTATTGCTGCTTTTAACTGTTATATCCCAAGCACCATTATACGCT 116
Db      100223 TGTCAAGAGCTCGATCATTCACCTCTCTTCACTAAGACATCATGATGCTCT 100169

RESULT 11
LOCUS      AE010854
DEFINITION Methanosarcina acetivorans str. C2A, section 199 of 534 of the
complete genome.
ACCESSION  AE010854 AE010299
VERSION     AE010854.1 GI:19915653
KEYWORDS
SOURCE      Methanosarcina acetivorans C2A
ORGANISM    Methanosarcina acetivorans C2A
            Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
            Methanosarcinaceae; Methanosarcina.
REFERENCE   1 (bases 1 to 11716)
AUTHORS    Galagan,J.E., Nussbaum,C., Roy,A., Endrizzi,M.G., Macdonald,P.,
            Fitzhugh,W., Calvo,S., Engels,R., Smirnov,S., Alnoor,D., Brown,A.,
            Allen,N., Naylor,J., Stange-Thomann,N., Dearellano,K., Johnson,R.,
            Limon,L., McKean,P., McKernan,K., Talamas,J., Tirelli,A., Ye,W.,
            Zimmer,A., Barber,R.D., Cann,I., Graham,D.E., Graham,D.A.,
            Guss,A., Hedderich,R., Ingram-Smith,C., Kuetner,C.H.,
            Krzycki,J.A., Leigh,J.A., Li,W., Liu,J., Mukhopadhyay,B.,
            Reeve,J.N., Smith,K., Springer,T.A., Umayam,L.A., White,O.,
            White,R.H., de Macario,E.C., Ferry,J.G., Jarrell,K.F., Jing,H.,
            Macario,A.J.L., Paulsen,I., Pritchett,M., Sowers,K.R.,
            Swanson,R.V., Zinder,S.H., Lander,E., Metcalf,W.W. and Birren,B.
            The Genome of M. acetivorans Reveals Extensive Metabolic and
            Physiological Diversity
            JOURNAL      Genome Res. 12 (4), 532-542 (2002)
TITLE      The Genome of M. acetivorans Reveals Extensive Metabolic and
            Physiological Diversity
PUBMED     21929760
REFERENCE   2 (bases 1 to 11716)
AUTHORS    Birren,B.
            Direct Submission
            Submitted (20-MAR-2002) Center for Genome Research, Whitehead
            Institute, Nine Cambridge Center, Cambridge, MA 02141, USA
FEATURES
source
1..11716
/organism="Methanosarcina acetivorans C2A"
/mol_type="genomic DNA"
/db_xref="taxon:188937"
complement(648..1085)
/gene="MA1763"
complement(648..1085)
/gene="MA1763"

```

```

/codon_start=1
/translation="transcriptional regulator, Copg family"
/product="transcriptional regulator, Copg family"
/protein_id="AA05169.1"
/db_xref="GI:19915654"
/translation="MSGDTMEAGLMRIGVSPDALLTFQDEIIKKRDSPPSSEAIRDA
LISYTYEMMEDIKGRVGTIAVIYHKKSLNSLINVGHSHLTKYSVMYLDT
DDCELIYDNGGEITELAGSIILAKVKRSKLITVDPNKII"
complement(2087..4807)
/gene="MA1764"
complement(2087..4807)
/gene="MA1764"
complement(2087..4807)
/codon_start=1
/translation="cell surface protein"
/product="cell surface protein"
/protein_id="AA05170.1"
/db_xref="GI:19915655"
/translation="MERFNMKLTTRKKPTGNLIRSLAVLLFSLVCTAAGEELYPEQ
EAVPSDTSVASLILYLDQAPSGLAGYGLNLISDPVAQITGVFFPMSGINSLSL
PADIVQIKAVDLERKINGDLDFLIGTIVTSIAGSEANLTVSVRLDDSENTISVT
VREKLTSGTISENPVVISLPBESTLSGASQVPEIRADKPTGEGISGVDFNLLENPA
VGEFTNVEYFPMVGLSENSMNPNTSIKRAVDINDVYKAGDENVTILAVTFAETRG
SEISLTFNRLLDDSGSNINAAVASVVOQSSTLPVAKFTATPTSGSPLTVOGTDE
STGSEPTAWTDFDGAATATQSPSHYTSAGNYTNLTVTNDGSDSEKTDYITVSE
SSTPEPEPVAFTADVTNGTVPLTVNFTDOSTEAPTSMAMPFDNDGTVSTEQNPST
TVTSAGCTVTVLTVANAGSDSEVKIDYITVSESTPEPEPVAFTADVTNGTVPLT
VNTDOSTGSPSHLMDPNDGNTSATQONPSHTNYAGNYTNLTVISGNSSEVKAD
YITVSESTPEPEPVAFTADVTNGTVPLTVNFTDOSTGMPSTNAMPDNDGNDST
EONPEYTYTAAGNYTNLTVSEVGSDESKVETVTDGSTTPARPDLVSSILPP
SSTPANTSIGNIGATISNTGESDAGTFNATLVNGAVDTHAVSGIVSGSATVNSKP
PVAAGYVAVKVSADAGNVIASDEENELADCIKYEBAKSGSOSGSSSVLNTVTV
PVCEPVPSPALKFGEISAGOTSEBOLLTLNKGVSIGIKYTAKEVDSMEDLFGGL
LDSQIMSDYCNVIGNSQENATLALHPSDVTGKNKGTITPAEVAE"
complement(4828..5748)
/gene="MA1765"
complement(4828..5748)
/gene="MA1765"
complement(4828..5748)
/codon_start=1
/translation="conserved hypothetical protein"
/product="conserved hypothetical protein"
/protein_id="AA05171.1"
/db_xref="GI:19915656"
/translation="MOKRSGVILISILAFATLLSALICTGLAEASPSLIIRPSEISL
GTGEICELSDNSIPBGLSGVYLTVLTPDPAVIOANFPMASLITVSELPASSV
KLKAVDLOALIGAEDEVELATLTKGEGSGTGTTVYSKRDDSENTIIIEGAEYP
KEDGDSSEDDGSDSEEDSEENTSEVENVTSSVKTSIDPLSOSNSSMSVIOQDP
NETAREMKSADENSLKKEVEISSPETSEKODDLIHSDADESKNDGTNNYTGNS
TGNSTLPGGFYISISFVFVIVYSRKK"
complement(5861..11104)
/gene="MA1766"
complement(5861..11104)
/gene="MA1766"
complement(5861..11104)
/codon_start=1
/translation="cell surface protein"
/product="cell surface protein"
/protein_id="AA05172.1"
/db_xref="GI:19915657"
/translation="MDPCYLIMWELIVHIALGFFLYGTGKAVTDRLIFVLLTLIC
SGVGLTFPTDAADAEWVGLPLTVQTEVTDLMFDTAPAPNMEQVVTIKFTLPE
AAVBERGIAVARIYASVCGMOSDYAFSINSMDGDGAYEQAQVPTGCAFOYI
IDPNWERTIGNDSEPDGCHGTQEPYIMIDHNRSYSDVFMVYDVLILYGTINAV
DTTGSDGRIKITYLVVANDSSTQTTYYVNEGDGCSYTEEDNREGAAVNTSFG
TGLGEVSLATLVDMASNNCCYGPETADNNPEYTGTPSVBGTFTTNQLDRDPTQ
GAYSGIDSNWNTSVYSGSDVILGARYFLAGTAAPYKIPLAFVKSPIETSSOP
VADFTADVTSGDAPLEVQFDASTGVSAYDFDNDGTVSTQONPMVTSYTEGYS
VULVTNEDGSDSELEKTDYITVTOAGQVATNDLSIGIVNTPASAVFARNTPEVUL
NYONQGTATLNIISIAVVASDVSSGVPTNTTIIAGDAKTATVLIPIITLDELG
TYYTTRAVTDPNLIAETDETNKNSASAPLAVYDYGKGGIYWEKGSNTTMMTFLQ
GNLVTSDPSAVQAVGESRTETWASLPDPDGSITIKAPLIVAYVYMDQPGGAPM
LNIINENGLINDGNISTGNGTLYRWSNGLAANBYGLCYVDVDTTPESAGNSLWVT
PVGENKALYPTSLVVVYGNDDNTRKQIINIEBCDGLSASSYGTTPBEATVAPPT
GISIDEKVTNMLYVSPAGSDPEGNLFPNGIIVATNMAQSGSNSSGSLAPATVYI
NATGNAGVQSTTSGMDALQOILVIEYESAEVAPSAVADFTATPTSGDAPLAVNFTD

```



```

repeat_region      5466..5853
                    /rpt_family="L2"
repeat_region      5884..6983
                    /rpt_family="L1"
repeat_region      7909..7936
                    /rpt_family="(TTTA)n"
repeat_region      7939..8944
                    /rpt_family="L1"
repeat_region      9538..9572
                    /rpt_family="AT_rich"
repeat_region      11515..11638
                    /rpt_family="MIR"
repeat_region      11820..12393
                    /rpt_family="L1"
repeat_region      12686..12756
                    /rpt_family="Alu"
repeat_region      12758..12888
                    /rpt_family="L1"
repeat_region      13182..13413
                    /rpt_family="L1"
repeat_region      18096..18137
                    /rpt_family="L1"
repeat_region      18540..18883
                    /rpt_family="AT_rich"
repeat_region      18954..19004
                    /rpt_family="ERV1"
repeat_region      19083..19203
                    /rpt_family="(CATATA)n"
repeat_region      19247..19288
                    /rpt_family="MER1_type"
repeat_region      19455..19488
                    /rpt_family="(CA)n"
repeat_region      19709..19755
                    /rpt_family="AT_rich"
repeat_region      20114..20167
                    /rpt_family="AT_rich"
repeat_region      20467..20833
                    /rpt_family="L2"
repeat_region      20876..21032
                    /rpt_family="MALR"
repeat_region      21612..21887
                    /rpt_family="L2"
repeat_region      21893..22075
                    /rpt_family="Alu"
repeat_region      22438..23791
                    /rpt_family="L1"
repeat_region      23792..24476
                    /rpt_family="L1"
misc_feature       24594..24960
                    /note="match to EST B1458889 (NID:g15249545)"
repeat_region      24660..24772
                    /rpt_family="MIR"
repeat_region      25939..26066
                    /rpt_family="MIR"
repeat_region      26474..26975
                    /rpt_family="MIR"
misc_feature       26615..26806
                    /note="match to EST B1458889 (NID:g15249545)"
repeat_region      27142..27187
                    /rpt_family="AT_rich"
repeat_region      27569..27894
                    /rpt_family="Alu"
repeat_region      29779..29801
                    /rpt_family="(TA)n"
repeat_region      29802..30127
                    /rpt_family="MER2_type"
repeat_region      30128..30184
                    /rpt_family="(TA)n"
repeat_region      30716..31118
                    /rpt_family="ERV1"
repeat_region      31162..31216
                    /rpt_family="ERV2"
repeat_region      31353..31388

```

```

repeat_region      31513..31702
                    /rpt_family="(TA)n"
repeat_region      31735..31758
                    /rpt_family="MER1_type"
repeat_region      31768..32055
                    /rpt_family="AT_rich"
repeat_region      32493..32513
                    /rpt_family="Alu"
repeat_region      33000..33140
                    /rpt_family="(TA)n"
repeat_region      33201..33512
                    /rpt_family="MIR"
repeat_region      33512..33512
                    /rpt_family="Alu"

Query Match      25.9%; Score 36; DB 9; Length 44532;
Best Local Similarity 70.6%; Pred. No. 4.8;
Matches 48; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY      39  CCCTCTTTCTCAGAAAGGGTGACTATTTGCTGCTTTATTAAGTATTCACCA 98
Db      13898  CTCTCTTTCTCTTATTAGTTAGTTAGCTAGCTGCTTTTATTTTTCATAA 13839

QY      99  AGCACCAT 106
Db      13838  ACCACCTT 13831

RESULT 13
AC103069/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-171118, *** SEQUENCING IN PROGRESS
AC103069
Rattus norvegicus
AC103069
VERSION
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Rattus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 245108)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Crete,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Davies,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,T., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,N., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,B., Hines,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
Hollins,B., Howells,S., Huylk,S., Hume,J., Johnson,B., Johnson,R., Jolivet,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kover,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lousubewa,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,B., Montemayor,J., Moore,S.,
Morgan,N., Morris,K., Morris,S., Muntlasea,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K.,

```



Pastermak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Ploner, F., Poindexter, A., Popovic, D., Prims, R., Pu, L., Pu, M., Putroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Rutz, S., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shartabeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Soza, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Vales, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, D., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhuesern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 245108)

REFERENCE  
JOURNAL  
TITLE  
AUTHORS  
Moriarty, K.C.

Direct Submission  
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 245108)

REFERENCE  
JOURNAL  
TITLE  
AUTHORS  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 13, 2003 this sequence version replaced gi:23268957.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Project name: GNDK  
Center project name: GNDK  
Center clone name: CH230-171118  
----- Summary Statistics  
Assembly program: Atlas 3.0  
Consensus quality: 210210 bases at least Q40  
Consensus quality: 214318 bases at least Q30  
Consensus quality: 217468 bases at least Q20  
Estimated insert size: 221660; sum-of-contigs estimation  
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

-----  
NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 4120: contig of 4120 bp in length  
\* 4121 4220: gap of unknown length  
\* 4221 230259: contig of 226039 bp in length

230260 230359: gap of unknown length  
\* 230360 233662: contig of 3303 bp in length  
\* 233663 233762: gap of unknown length  
\* 233763 235430: contig of 1668 bp in length  
\* 235431 235530: gap of unknown length  
\* 235531 238472: contig of 2942 bp in length  
\* 238473 238572: gap of unknown length  
\* 238573 241059: contig of 2487 bp in length  
\* 241060 241159: gap of unknown length  
\* 241160 245108: contig of 3949 bp in length.

FEATURES  
source  
Location/Qualifiers  
1..245108  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-171118"  
91407..93045  
/note="wgs\_contig"

misc\_feature  
misc\_feature  
misc\_feature

ORIGIN  
Query Match  
Best Local Similarity 64.6%; Score 35.6; DB 2; Length 245108;  
Matches 53; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 17 AACTTAATTAATGCTGCTCAACCTCTTTCTTCAGAAAGGCTGACTATTGCTCG 76  
Db 188425 AAAAAAAAAAAGCTTCTCTTACAGATTGTGTAAAGAAAAAGAAATATCTCTTG 188366

Qy 77 TTTATTAACTGTTTATCCCA 98  
Db 188365 TTTCTCACTGTTTTCACAA 188344

RESULT 14  
AC113887  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-330P6, WORKING DRAFT SEQUENCE, 3  
unordered pieces.  
AC113887  
AC113887.5 GI:25138317  
VERSION  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_PULLTOP.  
KEYWORDS  
Rattus norvegicus  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 162561)

REFERENCE  
AUTHORS  
Allen, D., Marie, Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Barnham, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.T., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, B., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idelbird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Lebow, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louised, H., Lozano, R.J., Lu, X., Ma, D., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,



Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Miosavljjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norrie, S., Nwankweli, O., Okmonu, G., Olarnungsoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Rellis, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, R., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Vals, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, K., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, O., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 162561)  
Worley, K.C.

Direct Submission  
Submitted (05-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 162561)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 20, 2002 this sequence version replaced gi:22856569. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

# ----- Genome Center Center: Baylor College of Medicine Center code: BCM Web site: <http://www.hgsc.bcm.tmc.edu/> Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Project name: GRNA  
Center clone name: CH230-330F6

----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 146385 bases at least Q40  
Consensus quality: 147074 bases at least Q30  
Consensus quality: 147633 bases at least Q20  
Estimated insert size: 150941; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length  
(see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved  
\* 1 159455: contig of 159455 bp in length  
\* 159456 159555: gap of unknown length  
\* 159556 160880: contig of 1325 bp in length  
\* 160881 160980: gap of unknown length  
\* 160981 162561: contig of 1581 bp in length.  
Location/Qualifiers  
1. 162561  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-330F6"

## ORIGIN

Query Match 25.5%; Score 35.4; DB 2; Length 162561;  
Best Local Similarity 57.8%; Pred. No. 5.7;  
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 30 CCTGCTTACCTCTTTTCTTCAAGAAAGGCTACTTTGTCTGTTATTACTGTT 89  
DB 85429 CCGGCCCAACCCCTTTTTCAGAGCTGGTGGCTGACCTCTGTAATTAAGTATT 85488  
QY 90 TATCCCAAGACCATATCAACGCTAGACTGTTCTTATTGTAACAC 138  
DB 85489 TTCAACAAACGCAACAAACCTTAGTATGTTTCTTAGATTAGTCCAC 85537

RESULT 15  
AC146720/c 126477 bp DNA linear PIN 01-OCT-2004  
LOCUS Medicago truncatula clone mth2-17n5, complete sequence.  
DEFINITION AC146720  
ACCESSION AC146720.23 GI:52782535  
VERSION HTG.  
KEYWORDS Medicago truncatula (barrel medic)  
SOURCE Medicago truncatula  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 126477)

AUTHORS Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B., Cook, D., Kim, D.  
and Roe, B.A.

TITLE Medicago truncatula BAC Clone mth2-17n5  
JOURNAL Unpublished

AUTHORS 2 (bases 1 to 126477)  
Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B., Cook, D., Kim, D.  
and Roe, B.A.

TITLE Direct Submission  
JOURNAL Submitted (23-SEP-2003) Department of Chemistry And Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

REFERENCE 3 (bases 1 to 126477)

AUTHORS Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B., Cook, D., Kim, D.  
and Roe, B.A.

TITLE Direct Submission  
JOURNAL Submitted (28-SEP-2004) Department Of Chemistry And Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

REFERENCE 4 (bases 1 to 126477)

AUTHORS Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B., Cook, D., Kim, D.  
and Roe, B.A.

TITLE Direct Submission  
JOURNAL Submitted (01-OCT-2004) Department Of Chemistry And Biochemistry,  
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

COMMENT ----- Genome Center  
Center: Department of Chemistry And Biochemistry  
The University of Oklahoma  
Center code: UOKNOR

FEATURES  
source Location/Qualifiers  
1.126477  
/organism="Medicago truncatula"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3880"  
/clone="mch2-17n5"  
/clone\_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 25.2%; Score 35; DB 8; Length 126477;  
Best Local Similarity 63.9%; Pred. No. 7.8;  
Matches 53; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 55 AAGAGGGGACTATTGCTGTTTATTACTGTTATCCCAAGACCATATATCAAG 114  
DB 118152 AAGAGGGGACTATTGCTGTTTATTACTGTTATCCCAAGACCATATATCAAG 118093  
QY 115 CTGACTGTTCTTATTGTTAACA 137  
DB 118092 TTATCTATATTCAAAAGACCA 118070

Search completed: August 3, 2005, 23:49:12  
Job time : 6699 secs